NUMERICAL AND STATISTICAL TIME SERIES ANALYSIS OF FETAL HEART RATE

M.A. Shariati

B.Sc., M.Sc.

A thesis submitted for the degree

of Doctor of Philosophy to the

Faculty of Science

University of Edinburgh

1992

ABSTRACT

In this thesis, numerical and statistical time series analysis techniques have been applied for the purpose of objective quantitative analysis of fetal heart rate (FHR) recording from the antepartum period (period of pregnancy before labour).

Currently FHR is routinely analyzed visually and the form of medical interpretations that medics make, are simply binary. Medical literature on fetal monitoring based on FHR, have been extensively searched and the current status of FHR and its components of interest to the medics for screening purposes have been presented coherently. Ambiguities and false expectations have been clarified. All the FHR components of interest used by the medics during the antepartum period have been addressed and analyzed numerically.

Prior to any successful analysis of non-stationary FHR, its baseline or trend (an important screening variable by itself) had to be estimated and removed from the original data in a statistically unbiased manner. This has been achieved via a first order bi-directional autoregressive filtering technique. Non-stationarities in both first and second moments had to be taken into account.

The detrimental effects of the baseline when analyzing FHR variability through one-dimensional scalar statistic namely the standard deviation have been studied and the need to use detrended data for this purpose has been emphasized. A new non-parametric approach to analyze the development of the FHR variability based on relative frequency histogram analysis has also been proposed.

A rule based routine has been devised which makes use of the rate of change of the unbiased estimated baseline and the original antepartum recorded FHR to numerically detect FHR accelerations. The accelerations are a very important screening component of FHR which are usually scanned for visually. This visual scanning technique is inefficient and inaccurate.

Stochastic time series analysis has been used to study the underlying random mechanism which gives rise to long term variability, the main component of FHR variability. The Box and Jenkins method of model identification and diagnostic checking was used on detrended FHR time series. Maximum likelihood estimation (MLE) in conjunction with Kalman filtering formed the parametric estimator.

Diagnostics performed on the identified model indicated the adequacy of a second order autoregressive model to parsimoniously represent the main dynamics observed in short, locally stationary blocks of the detrended FHR data. Hence, the current main obstacle in the clinical routines for objective quantification of long term FHR variability patterns is eliminated, and quantities not available for the researcher and clinician so far, are made available. The scheme may be viewed also as a means of data reduction of a highly redundant information source.

DECLARATION OF ORIGINALITY

This thesis and work reported within it was composed entirely by myself in the Department of Electrical Engineering at the University of Edinburgh, Scotland, UK.

M.A. Shariati

March 1992

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. James H. Dripps for his guidance, support and constant encouragement. Prof. Peter M. Grant for his support. Dr. Ken Boddy for his assistance in providing medical expertise. Dr. Bernard Mulgrew and Dr. Steven Mclaughlin for their support. Numerous individuals have made the duration of this project productive, Dr. Ahmad Mirzai, Dr. Hany. E. Bassil, and Dr. Mike Ansourian merit special mention.

Finally I especially wish to thank Dr. Hadi Shariati (University of Sistan and Baluchistan, Zahedan, Iran) for his encouragement and financial support.

Dedicated to both
my Mother and my Father
for their patience and support

Table of Contents

List of Abbreviations

FHR	Fetal Heart Rate
FHP	Fetal Heart Period
FHRV	Fetal Heart Rate Variability
BPM	Beats per minute
NST	Non-stress Test
SNR	Signal to Noise Ratio
ML	Maximum Likelihood
MLE	Maximum Likelihood Estimation
ACF	Theoretical Autocorrelation Function
ACVF	Theoretical Autocovariance Function
SAC	Sample Autocorrelation
SPAC	Sample Partial Autocorrelation
tSAC	t-statistic of Sample Autocorrelation
tSPAC	t-statistic of Sample Partial Autocorrelation
RSAC	Residuals Sample Autocorrelation
RSPAC	Residuals Sample Partial Autocorrelation
AR(p)	Autoregressive Model of order p
MA(q)	Moving Average Model of order q
ARMA	Autoregressive Moving Average Model
BAS	Estimated FHR Baseline
ROAD	Rate of change of the FHR Baseline
ms	Millisecond
Hz	Hertz
FFT	Fast Fourier Transform

CHAPTER 1

INTRODUCTION AND THESIS LAYOUT

1.1 General Introduction

Fetal Heart Rate (FHR) is one of the important components of fetal biophysical profile and is used in long term monitoring for the assessment of fetal well being during the antepartum period (period of pregnancy before labour). However, no successful method has been proposed to quantitatively represent and analyze the variety of random, non-white patterns seen in FHR. The general aim of this project is to address this issue and to bring new, meaningful and hopefully applicable ideas and statistics to provide objective numerical analysis of FHR time series.

Currently FHR is generally still analyzed visually and on a *subjective* basis. This is an imprecise approach and has been the main reason behind conflicting opinions and interpretations amongst different groups of experts when given the same FHR recordings (as has been demonstrated so many times). Even if their interpretations match, they are generally *binary*, i.e. they are looking for presence of certain patterns, with virtually no attempt being made to numerically measure and analyze the patterns and components of interest.

During the course of the study a good insight was gained regarding the FHR variable and its limits and capabilities as a monitoring tool. One important point which

emerged was that FHR should not be used as a *single parameter* on its own for making decisions on fetal well being. This insight was gained, from an extensive, medically oriented literature search on "fetal heart rate monitoring techniques". This search has resulted in a useful and relevant set of information, at the beginning of this thesis, on FHR and its main features (or components) of interest, its limits and capabilities.

Biomedical time series such as FHR, exhibit varying patterns in time. FHR is a *nonstationary* time series, exhibiting both nonstationarity in first and second moments. For example, the data may exhibit slow underlying rhythms or trends (time varying first moment), and continuously statistically significant changes in variability (time varying second moment; in another words the variability, itself, has a time-varying mean value) over the changing trend. Proper, unbiased estimation of FHR trend (or baseline) is of paramount importance before any successful numerical analysis can be made concerning FHR main patterns of interest. This problem is being addressed for the first time in this thesis. The estimated baseline itself is also an important medical screening tool.

Both classical, one-dimensional statistical analysis and model-based, stochastic time series analysis have been applied to numerically extract useful information about FHR variability (correlated fluctuations superimposed on the baseline) from short, locally stationary segments of detrended FHR time series.

From the classical, scalar routine the amount of variability (which is what the medical experts currently estimate on a visual basis) is correctly analyzed, due to the use of unbiased detrended data. However such one-dimensional statistical analysis fails to represent patterns observed in the quasi-stationary windows of observation.

Ideally a more complex and realistic approach must be taken, which recognizes that the fluctuations of the normal FHR, while apparently random (stochastic), are not patternless. The FHR time series presents a variety of patterns for which a one-parameter description is inadequate, and a model of higher complexity is required to adequately describe the data.

In short a multiparameter model is necessary to describe the FHR time series. In this study, univariate stochastic time series analysis has been applied to address this issue. This approach has led to a parsimonious representation of the stochastic mechanism responsible for generation of patterns observed in short locally stationary segments of detrended FHR time series in terms of the parameters of the chosen model.

The Box-Jenkins method of time series model identification was used to identify the parsimonious stochastic structure required to represent numerically (through its parameters) the long term FHR variability patterns.

Exact maximum likelihood estimation (MLE) implemented via a linear Kalman filtering structure formed the parametric estimator (the ultimate approach). The theory behind these routines and their application have been thoroughly and coherently presented in this thesis.

Finally, one of the benefits of applying time series analysis to model detrended FHR data, was the successful classification of long term FHR variability patterns via its spectral property. As a result, numerical detection of *pseudo –periodic* behaviour in detrended FHR time series was also made possible. This pseudo-periodic behaviour takes the form of random correlated cyclical patterns which have often been clinically observed and reported[1, 2].

1.2 Thesis Layout

Chapter two introduces the reader to the concept of FHR monitoring, explaining how FHR is used by doctors, what features in the FHR data are of interest to them and how FHR is captured. The following questions were considered. What are FHR limits as a monitoring variable? What are its statistical characteristics? Has it got any predictive powers in regards to fetal condition? Are there any other important physiological fetal parameters? Has there been any attempt made to attach medical significance, in regards to fetal condition, to variability patterns observed over the baseline? These were amongst many other questions which had to be answered in order to clear up misunderstandings, to find a sense of direction, and be able to set feasible and realistic objectives for the duration of this study.

In chapter three the problem of nonstationarity in the mean (baseline) was addressed, and the time varying mean was estimated efficiently using a linear bi-directional digital filtering technique which avoids phase shifts. Also the signal acquisition and processing which produced the averaged FHR time series was briefly explained in this chapter. Finally, the effect of baseline *wander* (slowly time varying mean) on the standard deviation statistic which is used to estimate the amount of statistical dispersion in long term variability, was investigated and discussed in this chapter.

In chapter four, a classical, scalar statistical approach was used to analyze long term FHR variability (using detrended data). This led to an original practical approach being devised and proposed which summarizes and shows accurately the statistical behaviour of FHR variability in antepartum recordings lasting several hours. A rule based routine is also introduced in this chapter which makes use of the rate of change of the unbiased estimated baseline and the original FHR time series to detect *accelerations*. These are a very important screening component of FHR which are normally assessed visually by doctors.

In chapter five, univariate time series analysis (Box-Jenkins identification strategy) has been applied for specification of a parsimonious structure to model long term FHR variability patterns observed in short, locally stationary segments of detrended FHR time series. The theory behind the identification strategy and its application has been presented coherently and comprehensively in this chapter.

Maximum likelihood Estimation in conjunction with Kalman filtering formed the parametric estimator used in this study and is the main topic of discussion in chapter six. Also Model diagnostics are presented in this chapter. This the third stage in time series analysis is where the adequacy of data fit (thus the identified model) is checked. The chapter ends with the presentation of a new method for classifying long term FHR variability in terms of the spectral properties of the FHR time series using the stability triangle of a second order difference structure.

Finally chapter seven presents a general summary and conclusions, and after some further discussion, ends with proposals for future work.

CHAPTER 2

FETAL HEART RATE MONITORING

2.1 Introduction

Before fetal heart rate (FHR) time series can be analyzed numerically, one must know how it is interpreted by the medics for the assessment of fetal condition. What patterns or features do they seek for their interpretation? Why is FHR a valuable screening tool and is FHR the only measured parameter of value for the purpose of monitoring fetal condition? These are some of the many questions which are considered here.

This chapter begins with a brief historical survey, followed by a general explanation of fetal physiological measurements and their relation to fetal heart rate (FHR). In order to understand the characteristics of the FHR time series, the methodologies used for capturing it, are discussed in detail. Finally, the vital FHR components of interest used for monitoring the condition of the fetus are introduced and explained in detail.

The benefits, and limitations of FHR as a monitoring tool have also been discussed and its position fully clarified. It is hoped through introducing all this relevant background information, a link will be made with the overall objectives of this project, which are, in summary, numerical representation of the important FHR patterns, for

the purpose of objective quantitative numerical analysis of fetal heart rate time series.

2.2 Historical Review

Monitoring the human fetus during its development in the uterus and its eventual time of delivery has provided the bioengineer and obstetricians with many challenges. Not only is the baby relatively inaccessible until a few hours before delivery but the few fetal signals which can be detected on the maternal body surface are extremely small and contaminated by maternally generated interference.

The presence of fetal heart sounds was officially reported first by *Francois Mayor* a Swiss surgeon in 1818[1]. The detection was made by simply placing his ear on the maternal abdomen, while attempting to hear the fetus splash about in its liquor.

Three years later *Lejumeau Kergaradec*, a French nobleman[1], apparently unaware of Mayor's report, again described the fetal heart tones. He suggested auscultation to be of value in the diagnosis of pregnancy, diagnosis of twins, and determining fetal lie and presentation.

As with many discoveries, the obstetricians of the time were slow to respond to Kergaradec's discoveries and recommendations. In 1833, Evory Kennedy of Dublin published an extensive book entitled *Observations on Obstetric Auscultation* [3].

At the turn of this century, the famous head or Pinard stethoscope; an ear trumpet like horn, became the first standard piece of equipment for use by general practitioners, midwives, and obstetricians for simple auditory phonocardiography.

Until the mid 1960's, the detection of fetal distress during labour relied solely on using the head stethoscope to detect gross changes in fetal heart rate and on the obser-

vation of fetal bowel movements in the amniotic fluid. Antenatal monitoring of the fetus was mainly consisted of listening for the fetal heart sounds; if present the fetus was still alive.

During the 1960's a revolution took place in the clinical management of labour. New electronic monitoring techniques enabled changes in fetal heart rate (FHR) to be observed in real time and correlated with maternal contractions during labour. This led to a reduction in fetal morbidity, but also introduced problems of its own, for example, an increased incidence of caesarian section deliveries.

Thus additional methods of monitoring such as fetal blood sampling and PH estimation were introduced to reduce the degree of unnecessary intervention.

During labour transducers can be attached to the fetus, simplifying the collection and analysis of biological signals. However antenatal monitoring, that is monitoring the fetus from conception to term 38 weeks later, relies on non-invasive measurements, and is a potential growth area for further research and introduction of new technologies.

Before discussing FHR measurement techniques it is useful to look in more detail at the fetus in its uterine environment and to assess its main medical problems.

2.3 Fetal Physiological Measurement

The various factors which influence the fetus and fetal heart rate (FHR) are shown diagrammatically in figure 2.1[4]. The figure clearly shows the complex nature of the interrelationships.

In early pregnancy uterine contractions are completely suppressed by progesterone

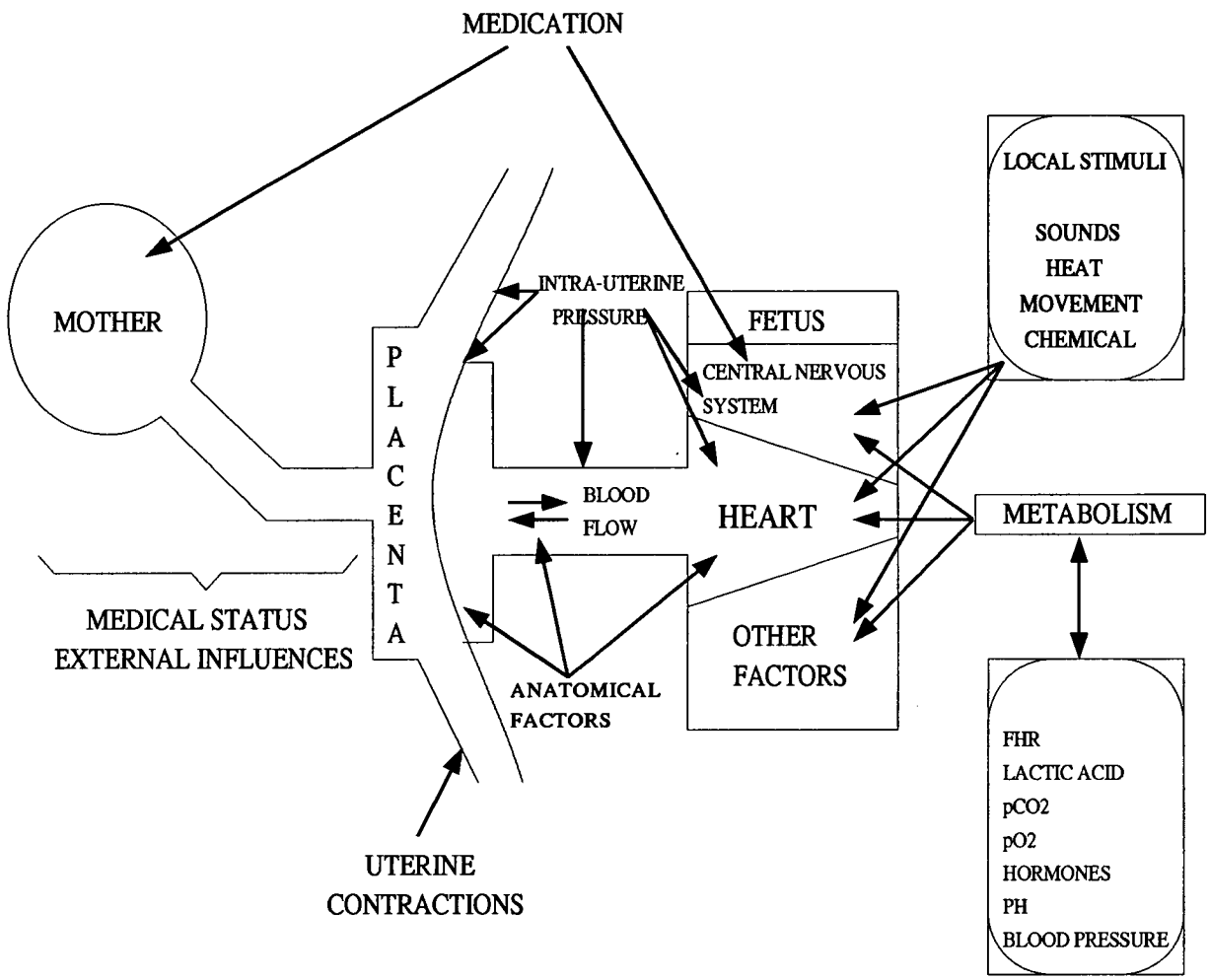


Figure 2.1: Simplified diagrammatical representation of the numerous factors which can interact to influence the condition of the fetus[4].

secreted from the placenta, but from approximately 24 weeks uterine contractility gradually reasserts itself. Contractions become coordinated and increase in intensity and frequency until labour is established[4].

As the intra-uterine pressure (IUP) increases during a contraction the maternal blood flow through the placenta decreases or stops completely, leading to a decrease in the transport of gases across the placenta and a resultant fall in fetal oxygen saturation. The available oxygen can also be influenced by a reduction in the functional capacity of the placenta caused by various chronic medical factors.

The cardiovascular status of the fetal heart, circulating hormones, the anatomy and position of the umbilical cord all have a role to play in determining the effectiveness of fetal blood flow and transfer of oxygen and waste materials to and from the fetus.

The response of the fetus to stress depends substantially upon its current PH, pO_2, pCO_2*, and lactic acid status. Lack of oxygen before, during or shortly after birth is the cause of at least 30% of the incidence of cerebral palsy and 10% of severe mental retardation [7].

Fetal asphyxia is a combination of hypoxia, hypercarbia and metabolic acidosis. Hypoxic metabolic acidosis is far more dangerous in the development of brain damage than respiratory acidosis[4].

Respiratory acidosis can be caused by prolonged or excessively frequent uterine contractions. In such a case exchange of CO_2 from the fetus is impaired leading to an increased pCO_2 and respiratory acidosis.

* pO_2 and pCO_2 stand for partial pressure of oxygen and carbon dioxide respectively[5, 6]. Partial pressure of either of these gases is a measure of their concentration in blood or tissue. The unit used for measurement is the Torr (1 Torr = 1 mm Hg at $0^\circ C$).

Anaerobic metabolism and lactic acid production is the result of a fall of pO_2 below normal levels. The production of lactic acid results in formation of additional carbon dioxide (Hypercarbia) and a fall in the blood PH. To differentiate between respiratory and metabolic acidosis requires both PH and pCO_2 to be measured, which is only possible with safety during the active labour phase.

The brain is a highly sensitive organ and has a very limited capacity to withstand prolonged hypoxic conditions. The brain-stem is more vulnerable than the cerebrum and when oxygen concentration levels fall below a safe level, disturbance in the control of respiration and temperature as well as disorders of muscle tone and reflexes may occur.

The fetus also responds to external stimulii such as movements, sounds and pressure variations with different types of changes in FHR. For example, external pressure on the fetal head usually produces a transient bradycardia (FHR below 120 beats per minute (BPM)) while compression of the thorax or abdomen results in a transient tachycardia (FHR above 160 BPM).

The medical status of the mother may also have an effect on the condition of the fetus and results in changes in FHR. The flow of blood to the intervillous area not only depends on the characteristics of the uterine contractions but also on other cardiovascular phenomena such as maternal blood pressure, haematocrit, oxygen saturation and maternal position. The use of drugs or medication by the mother also influences fetal condition. Even the mother smoking a cigarette causes a change in fetal behaviour.

2.4 Monitoring the Fetus

The main purpose of fetal monitoring is to separate the normal fetus from the compromised fetus so that appropriate medical intervention can be instigated. Thus

the problem is to find that variable or combination of variables providing the highest predictive values. Antenatal monitoring of the fetus provides information on fetal development, and may also be predictive of the fetus's ability to undergo stresses involved in birth.

Much work has been carried out to define a *single* parameter which can be used to identify the distressed fetus. However, most of the time the available systems require additional tests to be carried out to confirm the diagnosis.

For example, the most commonly used method to assess the fetal status, the FHR, can be affected by O_2 and CO_2 saturation, transport and exchange of blood gases in the intervillous space, maternal uterine contractions, fetal and maternal medical status and a host of other factors. Therefore, to confirm whether an event such as a rapid fall in the FHR is a normal transient response to uterine contractions or an indication that the fetus is hypoxic, one requires the fetal blood to be measured for PH, pO_2 and pCO_2.

Therefore it must be stressed that the use of FHR on its own should be regarded only as an important screening tool rather than a complete independent diagnostic tool. FHR is only one of a number of vital parameters used for monitoring the fetus[3, 4, 8] (see Fig. 2.1 and 2.2).

One conclusion with which most observers agree is that a normal FHR pattern, correlates well with normal cardiovascular function[9], with a very low false negative rate.

It should be emphasized that this does not guarantee a *normal* fetus at delivery as the fetus may be anatomically abnormal while having normal cardiovascular function. Although certain criteria for normality of FHR patterns have been established [2], there is a high percentage of false positive results (a baby in good condition at birth

following a suspicious or abnormal FHR pattern).

The signal sources, parameters measured and the information derived from the fetus and mother are shown in figure 2.2. It can be seen that the same basic information can be obtained by different methods. For example, the FHR can be derived from (1):- the signal obtained from the Doppler shift in an ultrasound beam, (2):- the fetal electrocardiogram obtained directly from the fetus (FECG) or indirectly from the maternal abdomen (AbdECG) and (3):- the fetal heart sounds (phonocardiography). Each of the methods has advantages and disadvantages which are dealt with in detail in the next section.

2.5 Detection of Fetal Heart Rate

There are two modes of FHR monitoring, *direct* and *indirect*.

The indirect or *external* mode utilizes Doppler ultrasound, phonocardiograms, or abdominal wall electrocardiograms to monitor the FHR. The direct or *internal* mode utilizes a fetal scalp electrode to monitor FHR through fetal electrocardiograms (FECG), and an intra-uterine catheter to monitor intra-uterine cavity pressure (maternal contractions).

The direct mode is restricted to the phase of active labor and delivery (*intrapartum* period) and the indirect modes are applied to the period of pregnancy before labour (*antepartum* period).

FHR monitoring with the direct FECG pickup at present is the method producing the best results in respect to *instantaneous* FHR recording. Instantaneous FHR is based on the interval between successive fetal heart beats (Fig. 2.3) and is a continuous $\frac{1}{I}$ plot where I is the interval[1, 3, 10].

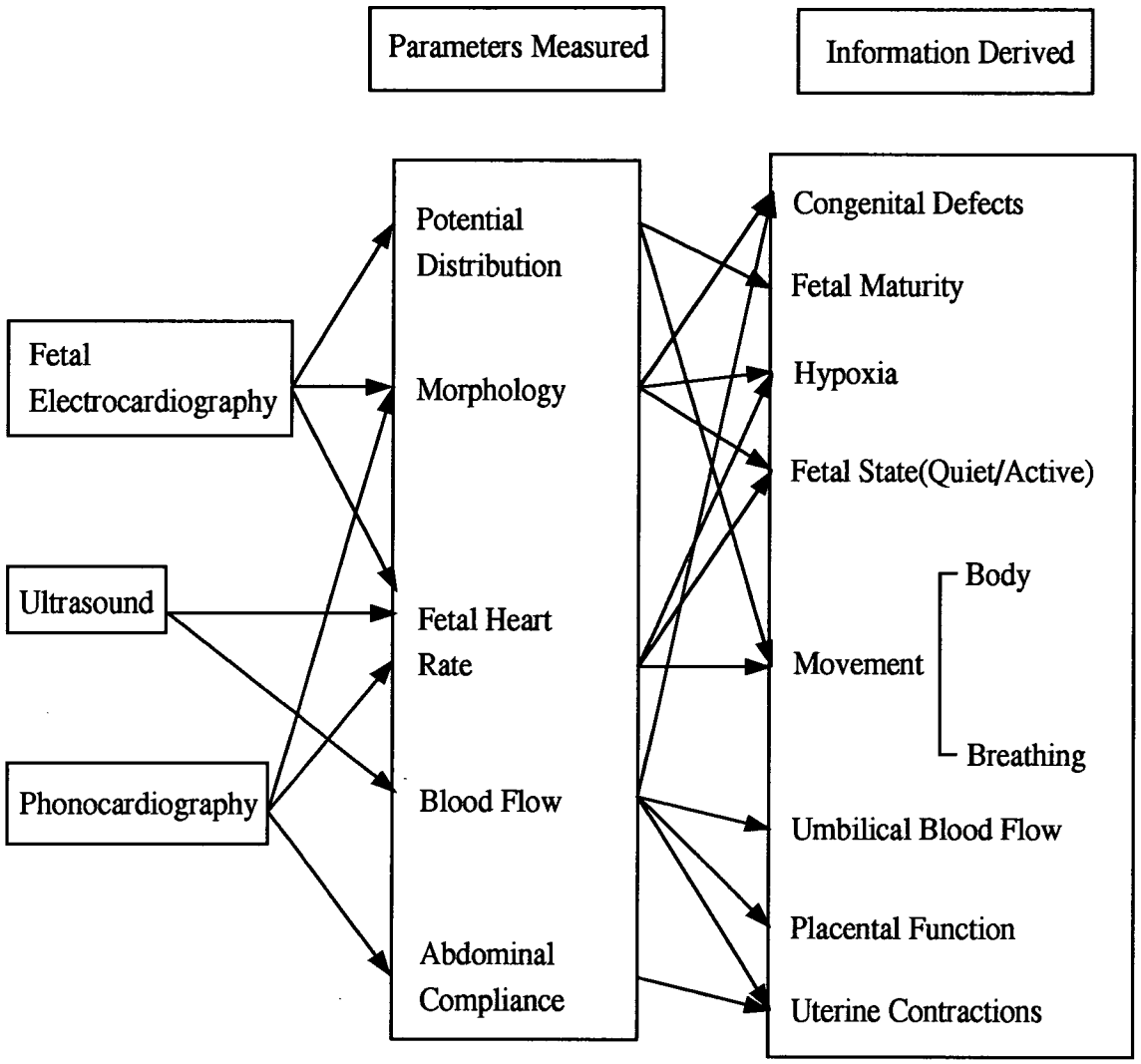


Figure 2.2: The signal sources, parameters measured and the information derived from the fetus and mother[4].

In the case of external modes, the signals used to derive FHR are invariably contaminated with noise, so to make the trace show a reasonable visual pattern, each FHR estimate is based on a one to two second block of measured data (i.e. it is averaged over several beats)[3].

Thus the term *averaged* FHR is more of an appropriate name to describe FHR calculated from external or indirect methods. As yet no indirect technique is presently available to compute FHR beat-to-beat variability, accurately and reliably for long term monitoring.

2.5.1 Electrocardiographic Derived Fetal Heart Rate

The electrical activity of the fetal heart can be detected at a number of different sites. The most common of these is the signal obtained by direct application of an electrode to the fetal scalp by piercing the fetal skin. The stainless steel electrode is either in the form of a double spiral or a hook with a reference electrode about 1 cm from the fetal electrode making contact with the cervix or vagina.

The FECG derived from this electrode pair has a QRS (see Fig. 2.3) amplitude of up to several hundred microvolts (100-200 μV). The electrode can only be applied once the membrane that fetus is contained in, has ruptured and the cervix had dilated by 1-2 cm. Once the FECG complex has been captured, the R-R interval can be measured. The interval is then converted to instantaneous FHR rate.

The obvious disadvantages of scalp FECG measurement are (1); its restriction to the phase of active labour, and (2); its invasive nature. Direct FECG pick up creates potential hazards such as infection and injury and due to its invasive nature, this method may also be rejected by patients for psychological reasons.

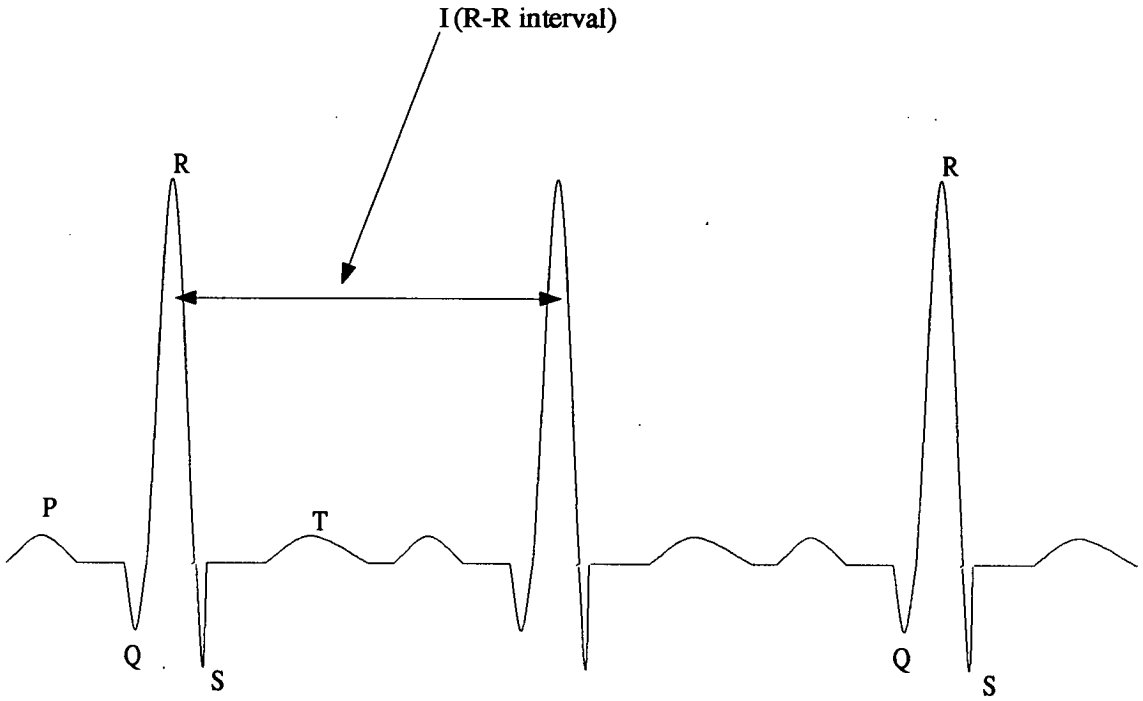


Figure 2.3: General structure of QRS complex which Scalp FECG will exhibit.

The FECG can also be detected indirectly from the surface of the maternal abdomen (AbdECG). The signal from this source is difficult to process, because it is contaminated by the maternal ECG and electromyogram (EMG). However, this is the only method of recording the FECG non-invasively during pregnancy.

A major disadvantage of AbdECG is when the small amplitude fetal signal is obscured when a maternal QRS complex is coincident with it. As a result, sophisticated signal processing is required to suppress the maternal complex as demonstrated by Frank et al and Callaerts et al[11, 12]. The methodologies used by both of these groups are not easy to apply and require expertise for their implementation. Also it seems their work is still at a preliminary stage and requires further tests and trials on patients, to verify its success rate as a reliable long-term external recording technique.

Finally, referring to earlier works, the accuracy of the indirect method of deriving ECG was investigated by Wheeler et al, Kariniemi et al and Carter et al. [13-15]. They all found that its success rate decreases significantly between 28 and 34 weeks, thereby limiting the value of the technique during this period.

2.5.2 Ultrasonic Detection of Fetal Heart Rate

The most common modality presently used for external FHR monitoring, both antepartum and intrapartum, is Doppler ultrasound[1, 16]. In this approach a transducer coupled to the maternal abdomen through a water based gel transmits at a frequency of between 2 and 5 MHz. A power level of less than 10 mW/cm^2 is used.

The transducer consists of one or more transmitting and receiving piezo-electric crystals producing a narrow or wide beam of pulsed or continuous wave ultrasound. Reflection of ultrasound occurs from discontinuities in refractive index within the beam. Moving discontinuities (such as heart walls and valves) cause Doppler shifted

reflections (backscatter) which are accepted through high-pass filters in the receiver to remove backscatter from stationary or slow moving discontinuities (e.g. the fetal chest wall).

The complexity and non-stationarity of this signal make it difficult to obtain beat-to-beat measurements of FHR, so averaging techniques are used by the fetal monitors to calculate the FHR from more than one pulse interval. Thus true beat-to-beat variability cannot be determined from ultrasound-derived signals.

The main advantage of ultrasonic monitoring is the ability to monitor throughout pregnancy from early in gestation. Also, with correct transducer positioning a high signal to noise ratio (SNR) signal can be obtained. However, with fetal movements, the heart can stray outside the beam resulting in total signal loss.

One of the main disadvantages of using Doppler ultrasound is that it is not passive, i.e., it transmits energy into the patient in a concentrated and directional manner[1,2]. Thus it poses a risk when used for long-term antepartum FHR monitoring. Medics have long questioned the safetiness of subjecting the mother and the fetus to long-term dosages of concentrated ultrasonic waves.

In summary it can be said that although ultrasonic techniques is an indirect monitoring method, it is not strictly non-invasive.

One way to obviate the fear of any risk from Doppler Ultrasound, is to use the abdominal fetal ECG. But, as has been mentioned earlier, it is not a reliable monitoring method for the whole antepartum period. As a result, this has led to renewed interest in the indirect Phonocardiographic methodology, which will be discussed next.

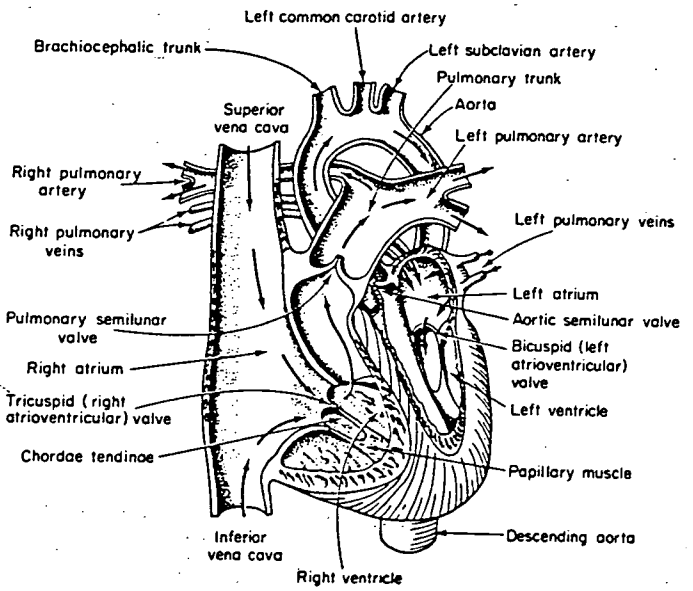
2.5.3 Phonocardiographic Derived Fetal Heart Rate

As the fetal heart beats, the closure of the heart valves produces sound that may be detected through mother's abdominal wall by listening with a stethoscope. This mechanical energy may be sensed by a specially designed microphone and amplified, producing an electrical signal that may then be reconverted to sound or used to produce an oscillographic tracing of the heart sounds known as a phonocardiogram. The amplified electrical signal can also be use as a counting source for an FHR monitor. Phonocardiography was indeed the first method used to record FHR electronically from the mother's abdominal wall[1].

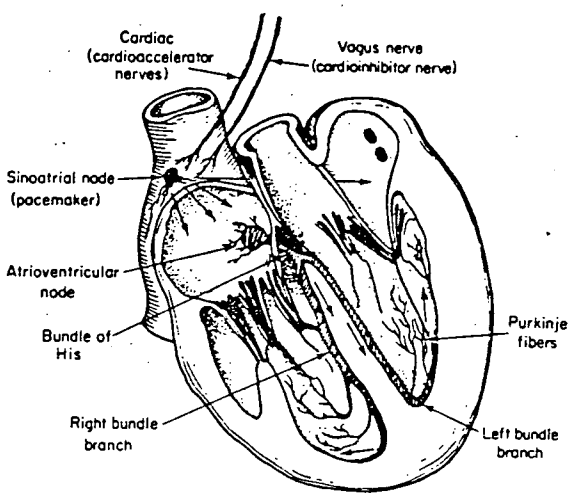
There are two major components to the signal. With ventricular systole (period of contraction of the heart muscle), the atrioventricular valves close, producing the first heart sound. With the onset of diastole (period of dilation of the heart cavities as they fill with blood), the aortic and pulmonary valves close, producing the second heart sound (see Fig. 2.4 showing the general heart structure). Given good positioning of the phono transducer on the abdomen of a suitable patient, the derived phonocardiograhic signal can result in less artefactual jitter than the Doppler ultrasound signal. This is due to the greater stationarity of the phono signal compared with the ultrasound signal which suffers from scintillation effects[1].

The main drawback to phonocardiographic FHR systems was that they were extremely sensitive to ambient noise such as maternal bowel sounds, movement of clothing, voices in the room, certain air conditioning systems, and especially noise produced by any motion of the microphone.

Recently however considerable improvements have been made in the development of a suitable acoustic transducer[17, 18], with a better signal to noise ratio. By maximizing the mechanical compliance of the transducer, attenuation of fetal heart sounds and



(a)



(b)

Figure 2.4: The Heart: (a): Internal structure; (b): Conduction system[5].

infrasounds are minimized. These new transducers are made using polyvinylidene fluoride membrane (PVDF)[19], or are inductive[20].

Because of the new transducers have a flat frequency response down to about 0.5 Hz, new information such as *fetal movement*, *fetal breathing*, changes in the systolic and diastolic time intervals, fetal maturity and function and strength of heart muscle[4] have been reported. The characteristics and patterns of fetal breathing movements detected by these new transducers may prove to be better as an indicator of fetal health than the presence or absence of movements as seen by real time ultrasound or detected by the mother.

In summary phonocardiography has the advantage of being non-invasive to both the mother and fetus. The procedure is also simple, so that it does not require the presence of an obstetrician. As a result, currently it makes the best candidate for long-term antepartum FHR monitoring. Indeed the antepartum FHR time series used in this project for conventional (classical) statistical analysis and parametric time series analysis, are derived from a phonocardiographic transducer.

2.6 The Evaluation of Fetal Heart Rate Patterns

The first important point to make is that FHR monitoring during the antepartum and intrapartum periods have to be considered independently, because the nature of the information is certainly different, and the informative parameters are probably different[8].

During the intrapartum period FHR information is observed along with the intrauterine pressure measurements, whereas during the antepartum period, the FHR trace is often the only source of information. It is important to mention here that, in the context of this thesis, we are only concerned with the antepartum period.

2.6.1 Antepartum Period

During the antepartum period which is the main period of pregnancy in terms of time, as part of a regular non-stress test (NST), obstetricians and midwifes visually scan the FHR recordings to detect normal variability in FHR, normal *steady state FHR baseline* (see figure 3.2) level, and a normal acceleration (for definition see section 2.7) count per hour.

The normal antepartum FHR characteristics [1-3] which also apply for the intrapartum period are:

1. A steady state baseline rate between 120-160 BPM (some prefer 155 BPM as upper limit[21]).
2. A peak to peak variability of greater than 5 BPM and less than 25 BPM.
3. An acceleration count of more than five in one hour of recording.

During pregnancy, *steady state FHR baseline* is normally in the range 120-160 BPM. Higher rates (tachycardia) can be caused by fetal hypoxia, by maternal fever, or by the effect of drugs. A rate lower than normal (bradycardia) along with decreased peak to peak variability ($<$5BPM) may indicate a dying fetus[2].

2.7 Accelerations

Accelerations of the fetal heart rate seem to occur most commonly in the antepartum period and in early labour[1, 3, 22]. It is important to mention that, the medical definition of FHR accelerations do not exactly correspond to the mathematical definition of the word acceleration. In general an FHR acceleration is defined to be a rise of FHR of at least 15 BPM above the *steady state FHR baseline* (where the baseline exhibits close to zero gradient activity) and maintenance of this rise for longer than 15

seconds, before falling back to a lower point where the baseline gradient will be zero [22]. After this completion point, the FHR might either follow a *steady state* trajectory, or rise up again for another FHR acceleration. Also refer to figure 3.2 for a pictorial representation of this phenomenon.

There are probably at least two physiologic mechanisms responsible for accelerations. Those seen in response to fetal movement or uterine contractions seem to have the same significance as fetal heart rate variability, i.e., their presence represents fetal alertness or arousal states. The other cause of accelerations seems to be partial umbilical cord occlusion.

The presence of fetal heart rate accelerations in the intrapartum period, as in the antepartum period, is reassuring. These accelerations may occur with contractions, fetal movement, or without apparent stimulus. The absence of FHR accelerations in the intrapartum period is not of itself alarming as long as variability is normal.

2.8 Fetal Heart Rate Variability

Present interest in FHR is based upon the idea that a given FHR pattern represents the output of a physiologic control system, which is the net result of the interaction of a number of opposing mechanisms[1, 3, 9, 10]. In the antepartum period and in the interval between uterine contractions in the intrapartum period, the interplay of these mechanisms results in unique correlated patterns, which are termed as fetal heart rate variability (FHRV).

These fluctuations are also held within close limits as the result of the interaction of cardio-accelerator and cardio-decelerator (or inhibitor) reflexes. These reflexes are usually in relative balance, so that there is a continuous superimposed irregularity which reflects the continuous hunting of the control mechanism. If, for some reason,

there is a transitory imbalance between these two opposing control mechanisms, it is reflected in a corresponding transitory FHR rise and fall (or vice versa) from a steady state FHR baseline.

In general there are two forms of variation, which are a combination of beat-to-beat changes and cyclic changes.

The beat-to-beat changes are commonly referred to as *short term variability*, and can only be detected accurately and consistently from scalp FECG, which is restricted to the intrapartum period (labour phase). Therefore short term variability can best be analyzed during the phase of active labour only and currently is subject to research to extent its availability to the antepartum period as well.

The random cyclic changes which are often referred to as *long term variability*, have been observed to exhibit periodicities of 8-30 seconds[1, 2]. Long term variability can be successfully detected and analyzed from proper indirect methods during the antepartum period.

As true beat-to-beat variation encompass only the components of highest frequency, which amount to less than one-tenth of the total FHR variation, Dawes et al [23] have questioned the value of beat-to-beat variation as an indicator of fetal well-being, because it is influenced by so many physiological variables.

They have come to the point of view that the medically most useful measures of FHR analysis are long-term variations and accelerations[23-25].

Furthermore, when obstetricians and midwifes visually analyze FHR variability, they base their judgement mainly on long term variations[3].

2.8.1 Factors Affecting Fetal Heart Rate Variability

One of the reasons why fetal heart rate monitoring has not fulfilled the expectations held at its introduction, may be the fact that many factors influence the firing frequency of the sinoatrial node (pace maker) and so complicate the reading and interpretation of the FHR patterns[26].

Among the factors which have influence on FHR and FHRV are maternal medication, fever, infection, body position, convulsions and shocks, and fetal age, behavioural states, brain damage and hypoxia. In this section the effect of fetal behavioural state, age, and hypoxia will be looked at in greater detail as these are the main contributory factors to FHR.

2.8.1.1 Fetal Behavioural State

The fetus, like the newborn infant, spends 95% of the time either in quiet sleep (30%) or active sleep (65%)[27]. Quiet sleep is characterized by a relatively stable baseline, presence of low variability (i.e. a peak to peak fluctuations of less than 5 BPM) and only occasionally short-lasting increases in heart rate during quick fetal movements (startles). According to Nijhuis et al[9, 28], a decrease in heart rate and a reduction in FHR variability (long-term variability) occur during quiet sleep.

During active sleep the fetus has a higher FHR baseline level and higher long term variability than the previous state, and accelerations occur periodically with fetal body movements (see Fig. 2.5 and 2.6).

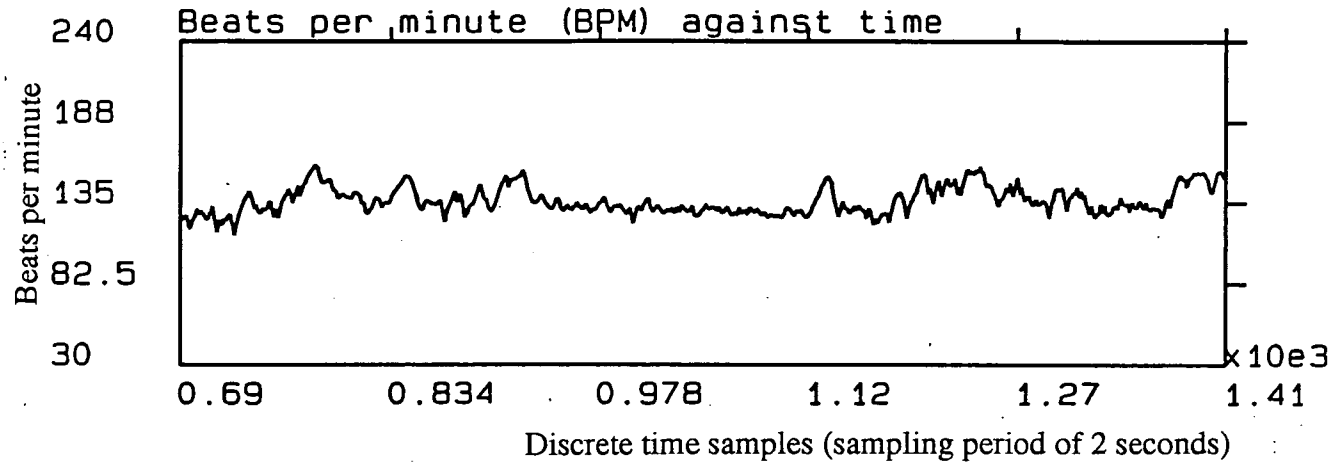


Figure 2.5: Part of an FHR tracing derived from an indirect phonocardiographic routine, showing a fetus which is probably mostly in active sleep state (this segment of recording represent a length of 24 minutes).

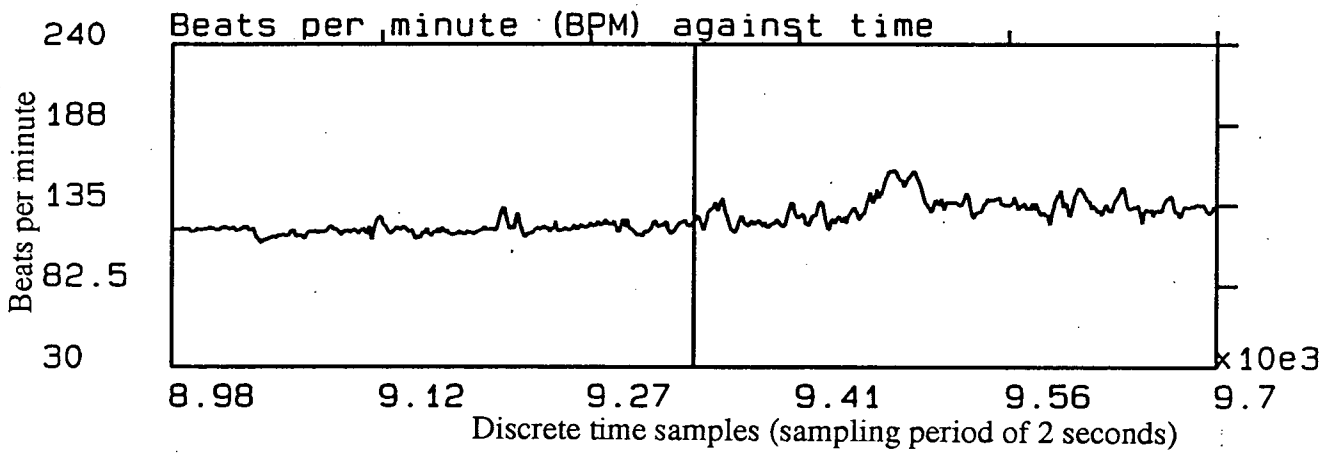


Figure 2.6: Another example, this time showing a probable transition from quiet sleep to active sleep (vertical line shows roughly the point of transition).

2.8.1.2 Fetal Age and Hypoxia

With increasing fetal age the following changes have been described in the fetal heart rate[26, 29]

1. Baseline heart rate decreases.

3. The incidence, duration and amplitude of accelerations increase.

4. Long-term variability increases in fetal heart rate.

The most ominous FHR pattern due to a hypoxic condition of the fetus is characterized by absence of accelerations, decreased long-term and short term variability. Earlier signs of fetal compromise have also been atributed to a significantly low blood *PH* and a transient increase in FHR variability.

2.8.2 Visual Measurement Methods of Fetal Heart Rate Variability

The visual appearance of FHRV was first described by Caldeyro-Barcia et al[30] in terms of irregularities of the basal rate. Hon[10] described a template, calculated for different heart rates, to place over the heart rate trace to measure the amplitude of these fluctuations as a percentage of the baseline rate.

Hammacher et al[31] described a value of less than five beats per minute amplitude of fluctuation (silent pattern) as abnormal, and values of between 5 and 10 beats per minute were regarded as the lower limit of normal. Variability above 25 beats per minute was also considered abnormal, and between 10 and 25 beats per minute amplitude, the variability was described as normal.

These values have been further acknowledged by subsequent studies[1, 32, 33] and are still considered the normal limits of visually discernible long-term variability.

Finally in a top international pilot study directed in Europe by van Geijn et al[34], they came to the conclusion that it seems trained obstetricians they have consulted evaluate heart rate variability by sight only and do not seem to attempt any form of calculations. This was detected by the fact that there were widely diverging opinion among the referees on the degree of heart rate variability.

2.8.3 Quantitative measurement of Fetal Heart Rate Variability

In general standard indices that have been used, reduce to some more or less intuitive estimation of standard deviation or variance. But these one-dimensional types of information extraction are not efficient and a more complicated, realistic approach needs to be devised. This approach should recognize that, FHR variability, while apparently random, is not patternless (see figure 2.5, and 3.5 to 3.10).

The FHR time series is a nonstationary time series, i.e, its statistical characteristics vary in time. This is a critical factor which must be considered, in order to have a meaningful quantitative analysis of FHR time series.

Non-parametric statistical estimation such as the standard deviation which might be reasonable in estimating the spread of FHR variability over the baseline, is not capable of quantitatively representing the non-white random cyclical patterns that long term FHR variability exhibits. Even when the standard deviation statistic is used, often it fails to represent the genuine dispersion of FHR over the baseline, because it has not been applied to the data properly (see the next chapter).

Therefore, there is the need to look at this problem and see if a parsimonious dynamic model can be found where, through the use of its parameters, one will be able to characterize different random cyclical patterns observed in long-term FHR variability.

2.9 The Biophysical Profile

Another very important antepartum test procedure which is valuable to be discussed here, is the fetal *biophysical profile test*, or simply the BPP test introduced in 1980 by Manning et al[35]. This approach is a commonly used and highly praised assessment routine which brings to attention the importance of other valuable variables which are used for checking fetal condition.

The profile consists of a combination of dynamic variables, (of which a test for *reactive – FHR* is one) combined, to produce a total BBP scoring. A fetal heart-rate pattern is considered *reactive* whenever there are two or more fetal heart-rate accelerations (refer to section 2.7 for definition) in every 30 minutes[3, 36].

The other variables which are always looked for first include, ultrasonic (or alternative) detection of fetal *breathing movements*, *gross body movement*, *fetal tone*, and *amniotic fluid volume* (*AFV*) (refer to Table 2.1[37]).

The BPP test starts with ultrasonic checking of body and breathing movements, and AFV score. If they show normal scorings, all is well. However, if any abnormality is detected, then an NST is carried out. This consists of passive FHR screening, looking for reactive FHR. Table 2.1 shows clearly the biophysical variables involved in a BPP scoring along with the technique and their interpretation.

In this procedure (as in other approaches), the observations made by experts on each variable seemed to be *binary*, that is it is present or not present, without any reference to duration or characteristics of the variable under observation.

Biophysical variable	Normal (score =2)	Abnormal (score = 0)
Fetal breathing movements	At least one episode of FBM of at least 30 sec duration in 30 min observation	Absent FBM or no episode of > 30 sec in 30 min
Gross body movement	At least three discrete body/ limb movements in 30 min (episodes of active continuous movement considered as single movement.)	Two or fewer episodes of body/limb movements in 30 min
Fetal tone	At least one episode of active extension with return to flexion of fetal limb(s) or trunk. Opening and closing of hand considered normal tone.	Either slow extension, with return to partial flexion, or movement of limb in full extension. Fetal movement absent
Reactive FHR	At least two episodes of FHR acceleration of >15 BPM and of at least 15 sec duration in 30 min(associated with fetal movement)	Less than two episodes of acceleration of FHR or acceleration of > 15 BPM in 30 min
Qualitative AFV	At least one pocket of AF that measures at least 1 cm in two perpendicular planes	Either no AF pockets or a pocket < 1 cm in two perpendicular planes

Key: FBM,fetal breathing movemement ; AFV,amniotic fluid volume; AF,amniotic fluid.

Table 2.1. Biophysical profile Scoring: technique and interpretation[37].

2.10 Conclusions and Discussion

Fetal heart rate variability (FHRV) derives from intrinsic beat-by-beat changes in stroke volume (short-term variability) as well as from longer-term fluctuations (long-term variability) resulting from the influence of the autonomic nervous system.

Short-term variability can only be detected accurately from FECG signals obtained from a scalp electrode (direct monitoring), which is restricted to the intrapartum period. Many attempts have been made without any success to compute short-term variation during the antepartum period accurately and reliably from indirect methods applied over a long period. Also, the value of short term variability itself as an indicator of fetal condition has been questioned, and its true medical significance is yet to be determined by medical researchers.

Medically during antepartum period most useful measures of FHR analysis are long-term variations, status of FHR baseline, and accelerations. Long term variability and accelerations are still assessed visually from graphic records of FHR.

Long-term FHR variability lends itself to numerical analysis because a major degree of data reduction is possible. Indeed in this project one of the main objectives has been to deal with this previously unexplored issue.

Some correlation exists between changes in long-term FHR variability and fetal condition. Normal variability has greater predictive value for fetal well-being than does reduced variability for fetal compromise.

The other very important FHR variable is accelerations. The presence of accelerations is used as a positive indication of fetal well-being. Again its detection is generally visual, which is an inaccurate way of scanning FHR data, for such an important variable. The presence of 5 accelerations or more in an hour of antepartum FHR trace is

a good indicator of fetal well-being. Again this was an important issue that this project addressed, i.e. detection of accelerations numerically via a computer.

It is important to state that fetal heart rate monitoring is only *one* of many monitoring tools. It is dangerous and misleading to claim that this technique can be used on its own for full evaluation of fetal well-being. It is more appropriate to refer to FHR as an important screening tool, and as an important component of the fetal biophysical profile.

In summary, fetal heart rate monitoring is certainly a useful screening tool for long-term monitoring of fetal well-being during the antepartum period. Its success as a continuous long term computerized screening tool is very much dependent on proper quantitative representation of FHR characteristics and patterns.

CHAPTER 3

BASELINE ESTIMATION OF FETAL HEART RATE

3.1 Introduction

During the antepartum period, as part of a non-stress test (NST), obstetricians visually screen the FHR data for normal patterns. There are, in general, three components in the trace which interest them; these are the position of the FHR baseline (or trend), occurrence of accelerations and the status of the variability.

Since FHR assessment is mostly visually based, their interpretation is unreliable as has been demonstrated by many prominent research groups such as Trimbos et al, Lotgering et al and van Geijn et al [34, 38, 39].

In comparison with simple visual assessment, objective numerical analysis of FHR, offers a better basis for understanding the physiology of FHR, and its changes under pathological conditions.

In this chapter, FHR time series obtained using an indirect phonocardiographic technique will be analyzed numerically to estimate baseline (or trend), and to study the effect of the latter on the standard deviation statistic which is the main statistical parameter for quantitative analysis of variability.

For successful detection of accelerations and for statistical and stochastic time series analysis of FHR variability, proper estimation of the baseline of the FHR time series is of paramount importance. Figure 3.1 shows the important position baseline estimation occupies in the overall structure, for the purpose of the objective numerical analysis of FHR. Figure 3.2 attempts to visually clarify what is meant by the FHR baseline, accelerations and long term variability.

But before starting our discussion on the above topics it is important to introduce the indirect monitoring technique which provided the FHR time series for our numerical analysis.

3.2 Indirect FHR Computation Methodology

As discussed in the previous chapter, indirect FHR computation using phonocardiographic signals is currently the best candidate for long-term antepartum FHR monitoring. Renewed interest in the phonocardiographic approach, is due to its non-invasive characteristics towards the mother and the fetus.

The FHR used for numerical analysis in this project has been provided from a phonocardiographic technique devised by Bassil[40]. Figure 3.3 shows the extent of signal processing involved in this technique, in order to successfully extract FHR information from noise corrupted fetal heart sound signals. Briefly these signal processing stages will be explained. The work reported in sections 3.2.1 to 3.2.3 inclusive was done by Bassil as part of his Ph.D. work.

3.2.1 Measurement Method

The heart sound signals are recorded using the Talbert and Southall TAPHO transducer[17, 18]. Typically, fetal heart beat signals had an amplitude of 10 mV while

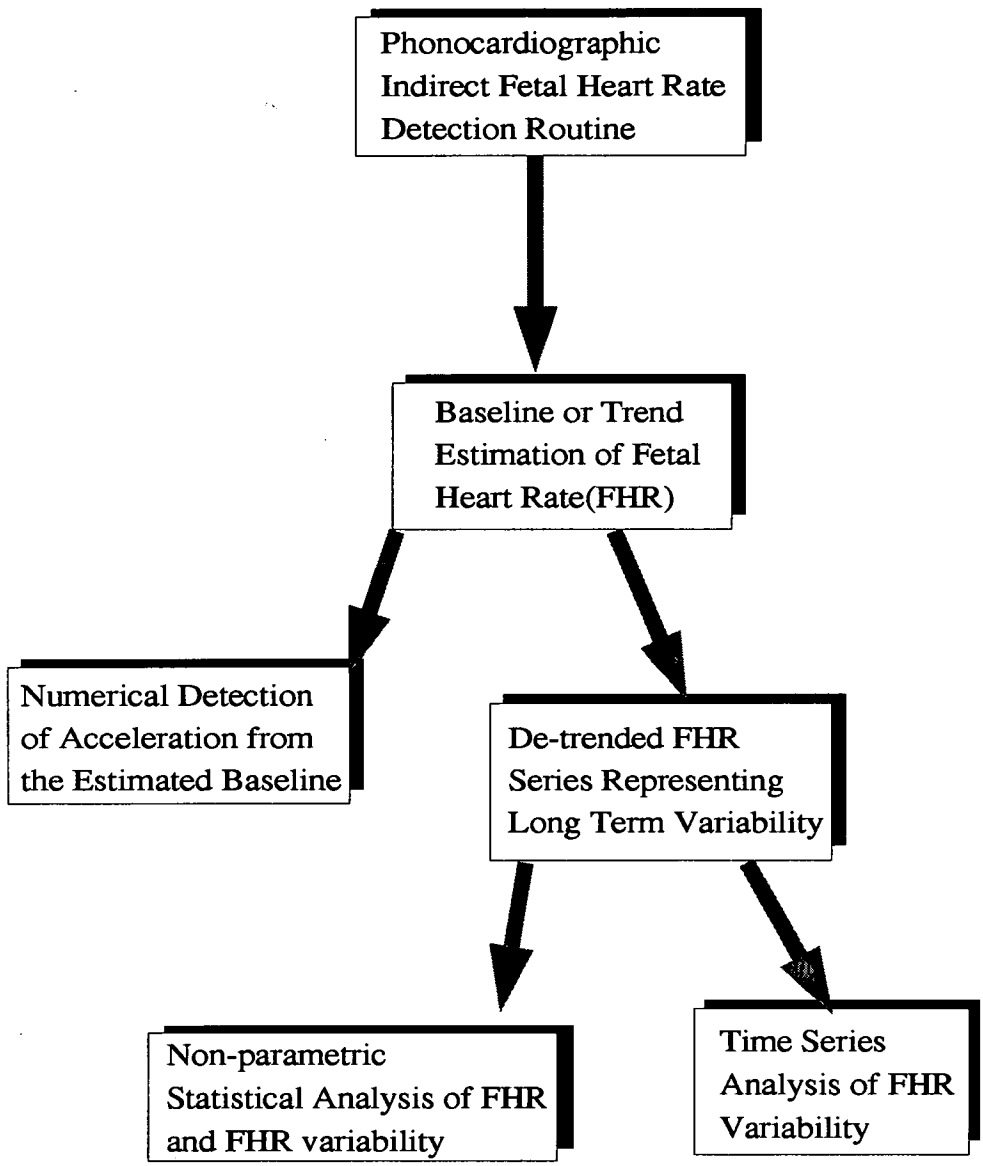
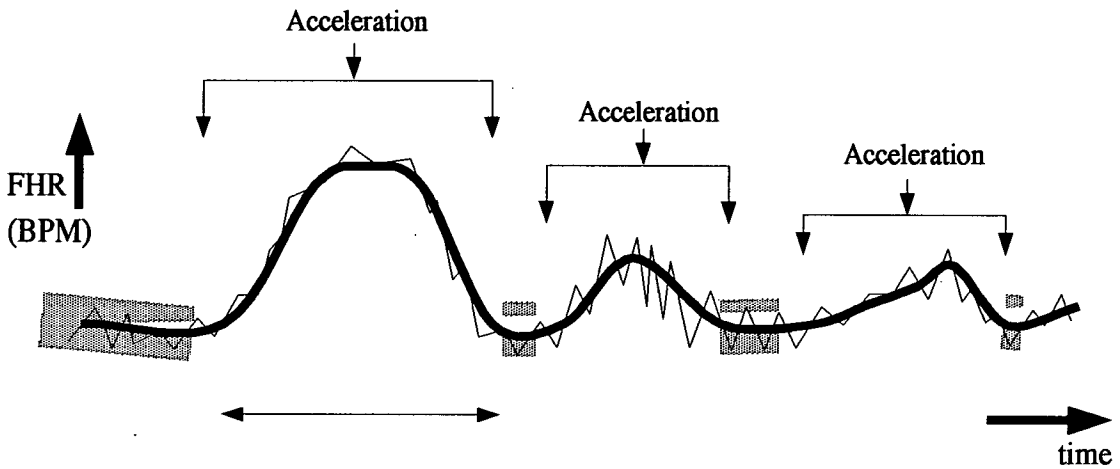


Figure 3.1: The general block structure for objective numerical analysis of fetal heart rate recording.



Typically an acceleration duration could be anything between 1 minute to 4 minutes.

——— :- Variability(Mainly Long-term FHRV in case of indirect measurement techniques)

━━━ :- Trend or Basline in FHR

▒▒▒ :- Steady state Baseline, where the rate of change in the estimated FHR trend is minimal

Figure 3.2: Illustration of FHR baseline and FHR variability and what is meant by an acceleration.

other interfering noises were about two orders of magnitude larger. Since phono transducers are susceptible to corruption by pressure waves from maternal and fetal movements, the patients are monitored during their resting period. The signal is recorded first on a standard Racal frequency modulation tape recorder without any analogue conditioning. This was essential, especially during the period of algorithm development, where replay of the same data was necessary and analogue pre-processing techniques were being optimized.

3.2.2 Analogue Pre-Processing

The phono transducer output signal was conditioned by being passed through a series of analogue circuits (Fig. 3.3). In order to enhance the signal to noise ratio of the fetal heart sound, it is first filtered by a 45 to 65 Hz band-pass filter.

The output of the band pass filter is then passed through a full wave rectifier, followed by a 20 Hz low-pass filter thus reconstructing the heart beat pulses. The filtered signal is then sampled at 50 Hz and divided into 2 second blocks (100 samples) for discrete conventional spectrum analysis.

3.2.3 Digital Signal Processing

The 2 second (100 samples) blocks are then zero padded to 256 points and their fast Fourier transform (FFT) is taken. Unfortunately, the global peak in the power spectrum of the 2 second block will not necessarily lie at the fundamental fourier component position (the heart rate). In order to avoid the real possibility of sub-harmonic errors being made by the FHR estimator, a further transform is performed on the linear power spectrum which leads to a global peak at a frequency corresponding to

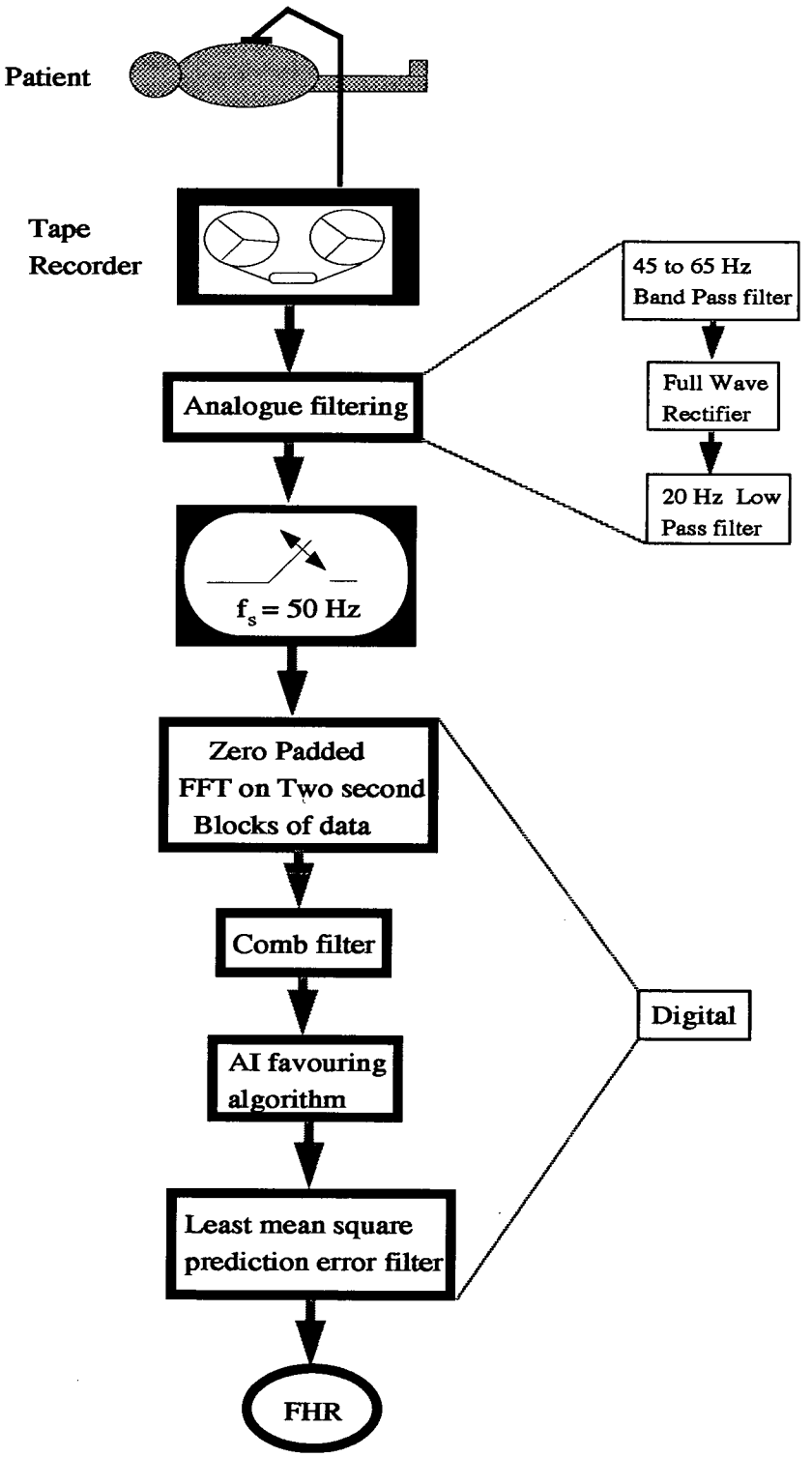


Figure 3.3: Processing Stages of phonocardiographic fetal heart rate estimation technique; AI: artificial intelligence; FFT: fast fourier transform.

the FHR [40]. Finally, an iterative interpolation technique based on successive interval bisection type search for the global peak is performed. Interpolation to the equivalent of a 4096 point zero padded FFT requires access to a normalised 4096 point $\frac{sinx}{x}$ (or sinc) function[40]. The above technique gives identical results to those obtained from a 4096 point zero-padded FFT but is much more computationally efficient.

The objective is to estimate as close as possible the position of the fundamental harmonic and hence the FHR. However, determining the FHR from a single peak can lead to a biased estimate in the presence of coloured noise. Also the global peak might not be the fundamental harmonic (it is often at the second harmonic position). Therefore a routine was used which estimated the position of the fundamental harmonic from a span of harmonics. This has been achieved using a comb filter [40, 41].

The global peak of the comb filter output is then expected to give a good estimate of FHR in the 2 second sampled block under observation. However, in some cases, close multiple peaks were observed due to asynchronous periodic interference and an artificial intelligence (AI) peak selection routine was devised to detect the most likely FHR peak[40].

Finally, the randomly distributed outliers which are still present at the output of the AI algorithm are removed using a least mean square (LMS) prediction-error filter.

The performance of the technique has been assessed by application to the ultrasound returns from a patient during the intrapartum period, allowing comparison to be made with instantaneous FHR (chapter 2) extracted from a fetal scalp ECG. A resulting error rate of approximately 3% was observed[40].

In summary, this external approach can reliably extract averaged FHR from long-term phonocardiographic recordings during the antepartum period (see part (a) of figures 3.5 to 3.10 as examples). The resultant averaged FHR time series has a sampling interval of 2 seconds and an averaging window of 2 seconds.

Due to the relative short size of averaging window, long term variability which is the main component of FHR variability (see chapter two) is fully captured. This is, in general, the main component used for both visual and numerical analysis of variability in FHR.

3.3 Baseline Estimation of Fetal Heart Rate

The estimation of baseline or trend is of fundamental importance in both visual and numerical analysis of FHR recordings. The baseline is an imaginary line representing a form of running average of heart rate and its proper unbiased estimation is an essential prerequisite for objective numerical analysis of fetal heart rate recordings.

This *time variation* in the trend of FHR constitutes a nonstationarity of the FHR process in the first moment or mean. Estimation and removal of mean nonstationarity from a time series such as FHR can be approached in two ways.

The first approach is to fit a deterministic model structure[42], such as a linear or non-linear polynomial structure, to short segments of data using standard regression analysis techniques. The other alternative is to apply some form of low-pass discrete filtering to separate low frequency components of the time series related to the trend from higher frequency components.

In regard to the first approach, a major problem arises in the case of long duration, mean-nonstationary time series such as FHR since the trend continuously changes in

shape (evolves) from block to block. Thus different deterministic model structures may be required from block to block in order to detrend the segmented data.

Clearly this first approach would be a very inefficient method of estimating the trend, especially when one is dealing with long monitoring sessions as in antepartum FHR monitoring. If applied to FHR, the curve fitting technique will result in unpredictable, and non-quantifiable statistical bias both in the estimated trend and in the detrended correlated series (original - trend) which represents the important, long term variability component of FHR.

The other alternative would be to estimate and remove the trend from the time series by passing the latter through a suitable linear discrete filtering structure. This is the approach which was taken in order to avoid the possibility of changing the statistical, frequency and phase properties of the estimated trend, and of the detrended data. The latter is used for FHR variability analysis.

The filtering technique used for baseline estimation/removal is based on a bi-directional first order autoregressive filtering structure, which was successfully applied by Orr and Hoffman in 1974[43] to detrend adult human heart rate. In 1983, this technique was also successfully applied by Dalton and Dawson [44] to averaged ultrasonically detected FHR. Historically, this trend removal filtering algorithm was introduced by Tukey[45] in a set of his published lecture notes in 1963.

3.3.1 Linear Estimation Method

In many filtering problems, one would ideally like the phase characteristics to be zero or linear. For causal filters it is impossible to have zero phase. However, for many digital filtering applications, it is not necessary that the unit sample response of the filter to be zero for $k < 0$, if the processing is not to be carried out in real time. One

technique used in digital filtering when the data to be filtered are of finite duration (N samples) and stored, is to process the data forward and then backward through the same stable filter. Such bi-directional processing effectively eliminates phase shift[43-46].

In regard to the problem here, the trend (or baseline) in fetal heart rate was calculated by a first order purely autoregressive low pass filter structure applied to FHR data first in the forward direction and then in the reverse direction. Equations (3.1) to (3.4), illustrate the general processing strategy:

$$y(k) = \rho y(k-1) + (1-\rho)x(k), \quad \text{for } k = 0,\ldots,N, \tag{3.1}$$

$$yr(k) = y(N-k), \quad \text{for } k = 0,\ldots,N, \tag{3.2}$$

$$zr(k) = \rho zr(k-1) + (1-\rho)yr(k), \quad \text{for } k = 0,\ldots,N, \tag{3.3}$$

$$z(k) = zr(N-k), \quad \text{for } k = 0,\ldots,N. \tag{3.4}$$

Also an additional final equation, (3.5) was added to make the overall real valued filter transfer function, (3.6) simpler[43],

$$u(k) = z(k) - (1-b)x(k), \quad \text{for } k = 0,\ldots,N, \tag{3.5}$$

where

$$b = \frac{2\rho}{(1+\rho^2)}, \text{ and } 0.70 < \rho < 0.95.$$

Equation (3.1) represents the time forward first order autoregression on the input FHR time series x(k), which yields the output series y(k). The output y(k) is then reversed in discrete time order using equation (3.2), allowing the data to be processed effectively backward via equation (3.3). This equation represents the same stable first order autoregressive structure, which has defined equation (3.1). Finally equation

(3.4) corrects the time order of final filtered output, to be in direct one to one correspondence with the original raw FHR data.

Both z(k) (equation (3.4)) and u(k) (equation (3.5)) can represent the underlying nonstationary trend of the original time series. u(k) is usually preferred over z(k), due to the fact that it makes the overall transfer function simpler.

It can be shown that the general overall impulse response resulting from the filtering steps of equations (3.1) to (3.4) (shown in figure 3.4(a)), is non-causal and has even symmetry. This is the reason behind the zero phase characteristics of this type of filtering. An alternative high order non-recursive FIR filter can also be used to implement the same filtering for trend estimation. This is achieved by using the truncated impulse response (for a pole at $z = 0.9$, inclusion of all h(k) with magnitude greater than or equal to 0.1 % of h(0) will be sufficient) shown in figure 3.4(a) to define the taps of this FIR filter.

Taking the *bilateral* (or *two−sided*) *z−transform* [46] of this symmetric impulse response, and setting z to e^{jwT} will result in a purely real valued open form infinite series transfer function, proving there is no overall phase shift. Alternatively through applications of z-transform theorems and their properties, to the discrete equations (3.1) to (3.5), a simpler real valued closed form transfer function can also be derived for this overall bi-directional filtering procedure, which is

$$H_{xu}(w) = (1-b)\frac{b\cos(w)}{1-b\cos(w)}, \quad where\,(0 \leq w \leq \pi). \tag{3.6}$$

The value of ρ in equations (3.1) and (3.3) determines the 3 dB cut-off point of this bi-directional autoregressive filter, as shown in the frequency magnitude plot of figure 3.4(b). Notice as ρ decreases in value the 3 dB cut-off point increases, effectively increasing the bandwidth of the bi-directional filter.

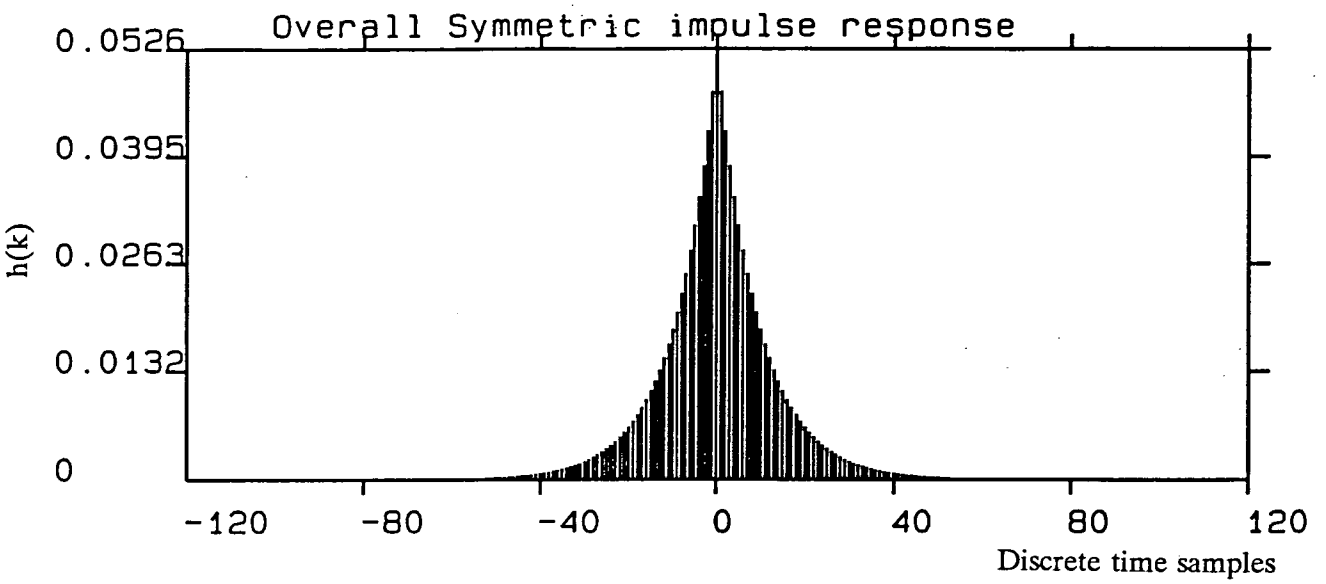


Figure 3.4: Part(a): Overall Symmetric, non-causal impulse response of the bi-directional filter defined by the equations (3.1) to (3.4).

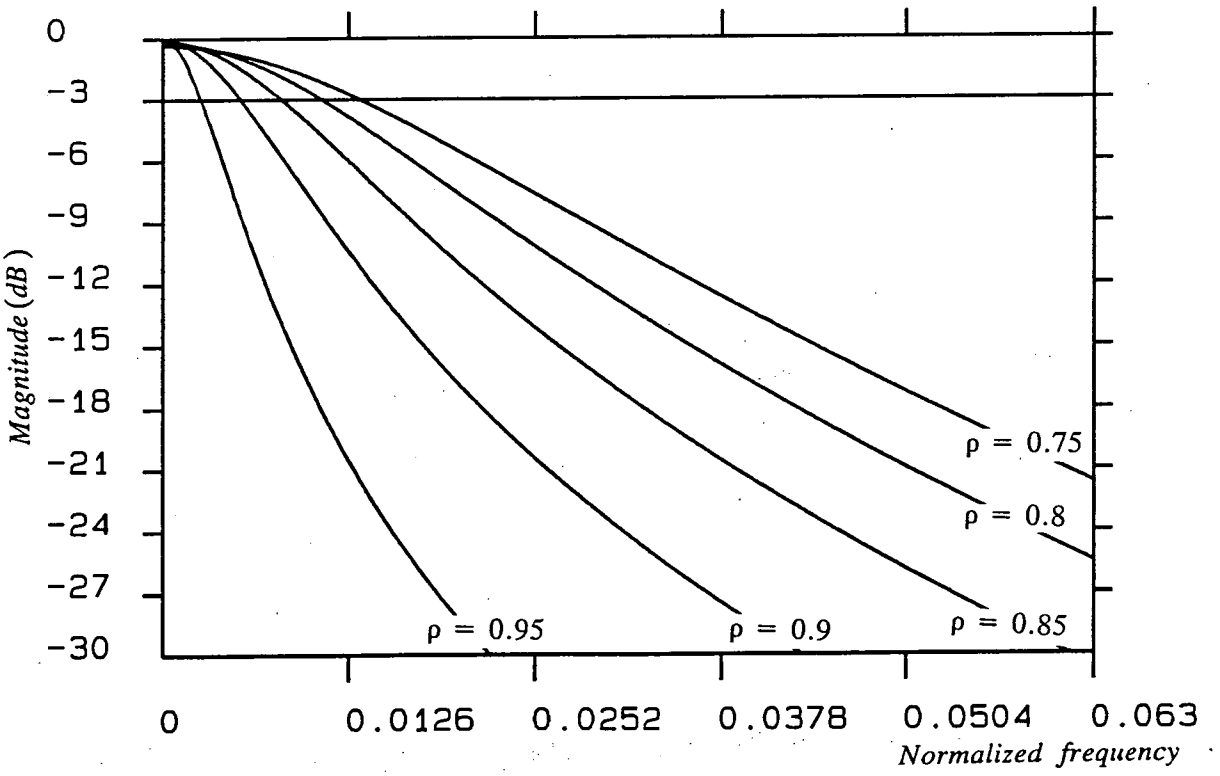


Figure 3.4: Part(b): Transfer function magnitude (in dB) against Normalized frequency of the first order autoregressive bi-directional filter, exhibiting different trajectories corresponding to different ρ values ($0.75 < \rho < 0.95$).

A value for ρ must be chosen such that the detrended FHR series will not exhibit any non-stationarity in the form of a pronounced trend which results in constant variation of the overall mean of the filtered series. Through numerous trials on phonocardiographic derived, averaged FHR, a value of 0.9 was found appropriate for ρ. Using this value resulted in a good estimate of the overall baseline of the FHR data which also included the long-lasting accelerations as part of the estimated trend. This property is exploited in the next chapter for the detection of accelerations.

The detrended FHR series is obtained as coloured residuals by subtracting the smooth trend, u(k), from the original time series x(k). This detrended series can then be used to analyze, quantitatively, the FHR variability (FHRV; see previous chapter). This may be done by applying either conventional scalar statistical analysis techniques or by applying much more numerically complex stochastic time series analysis. Figures 3.5 to 3.10 show how six 24 minute blocks of FHR (each corresponding to 720 samples) have been detrended, acquiring stationarity in the first moment.

Finally, parts (d) of figures 3.5 to 3.7 show FFT based power spectra taken from the original FHR data and from their corresponding detrended versions shown in figures 3.5 to 3.7 (parts (a) and (c)). These show how much the power in the low frequencies has been reduced by the trend removal operation.

All the power spectra shown in parts (d) of figures 3.5 to 3.7 were based on the FFT of 720 samples of real data zero-padded to 1024 points. A Hamming tapered window was used to reduce leakage[42, 47, 48].

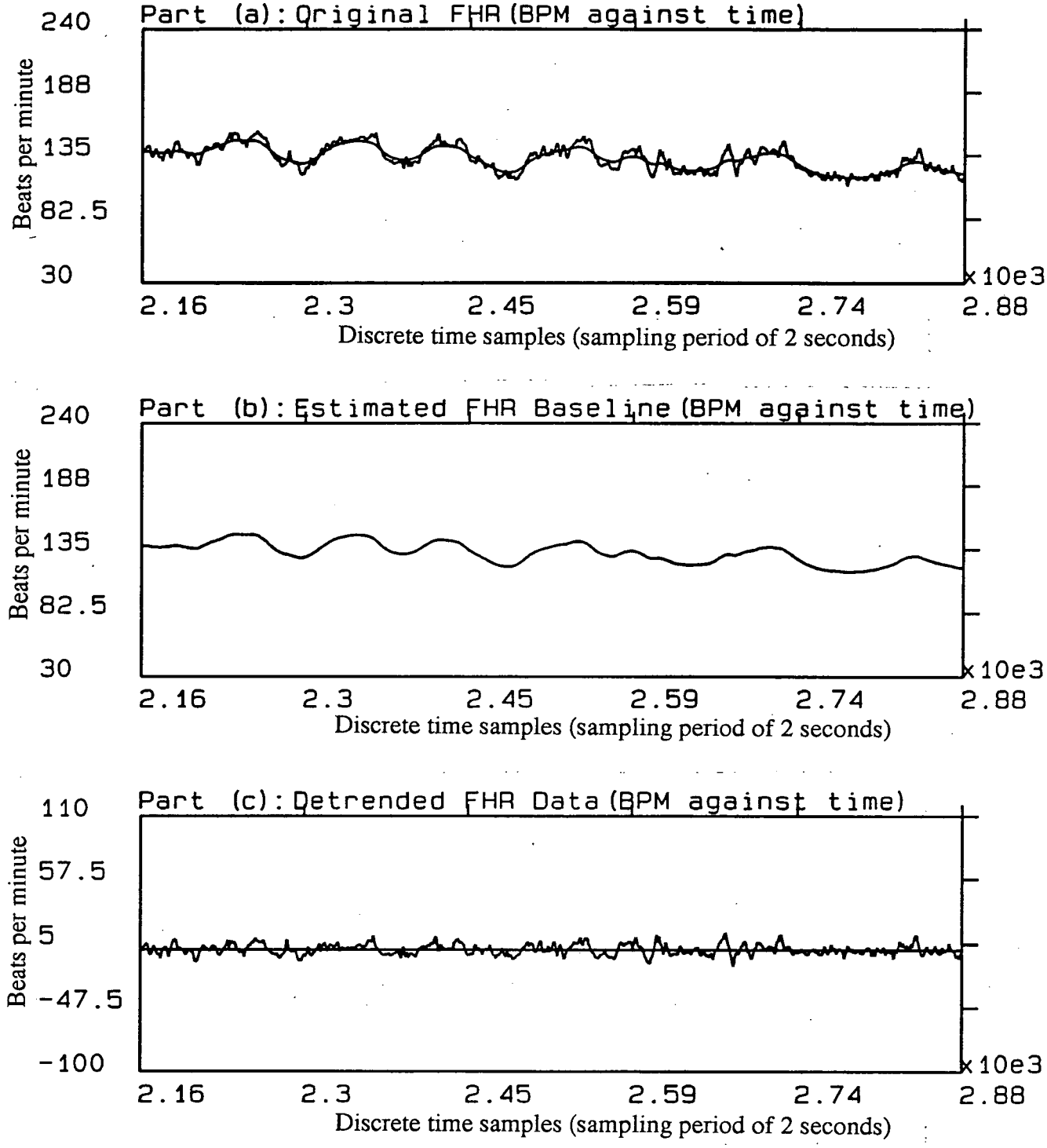
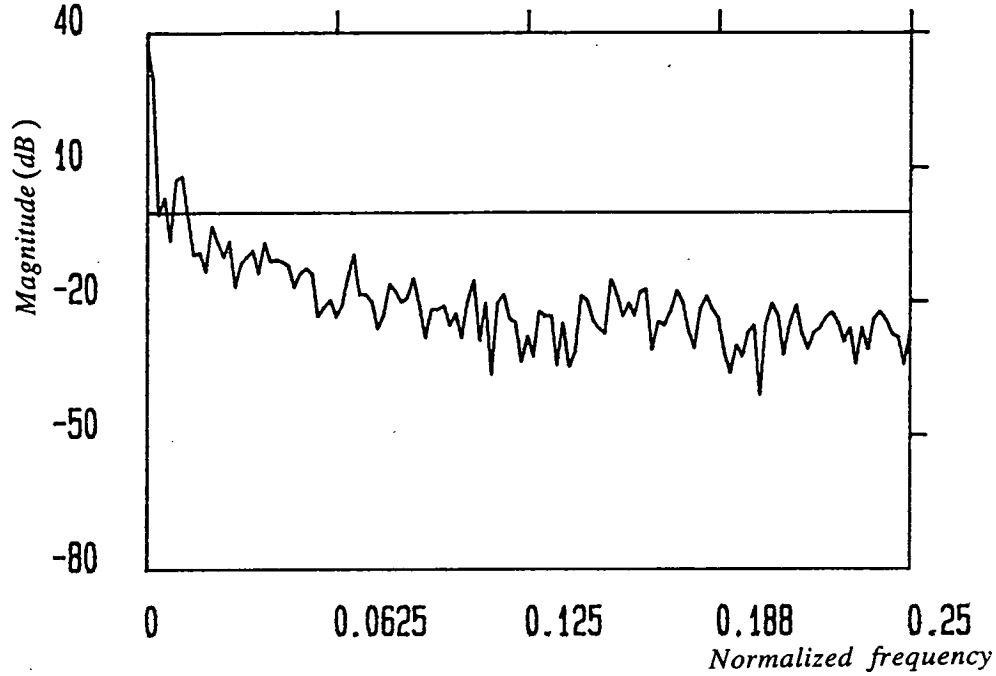


Figure 3.5: Part (a): Presents 24 minutes of averaged phonocardiographic FHR data, along with its estimated baseline.
Part (b): Shows the estimated baseline on its own.
Part (c): Shows detrended, mean stationary, averaged FHR data.

Part (d): FFT based power spectrum from the original FHR data.



Part (d): FFT based power spectrum from the detrended version.

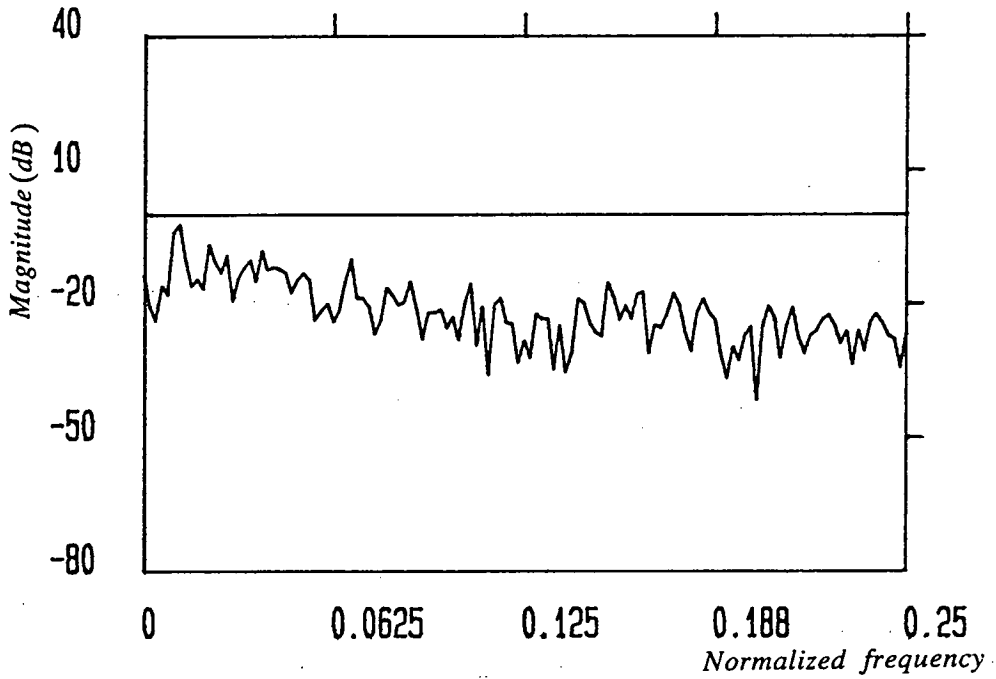
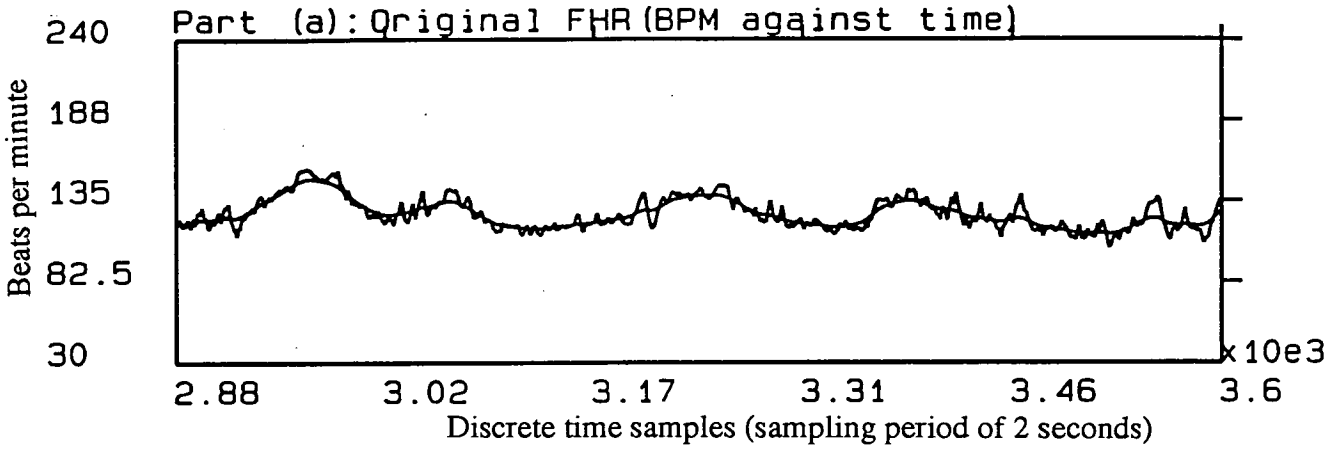
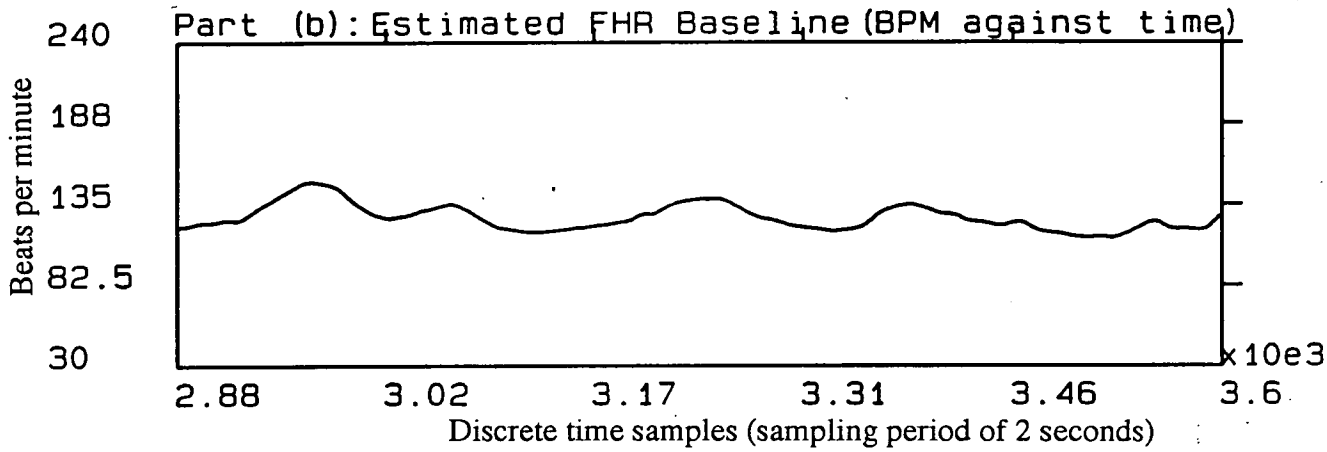


Figure 3.5: Part (d): FFT based power spectrum taken from the original 24 minutes FHR data segment (part (a)) and from its corresponding detrended version (part (c)), showing the effect of the trend removal operation on the spectral content of the FHR data (Magnitude (in dB) against normalized frequency).





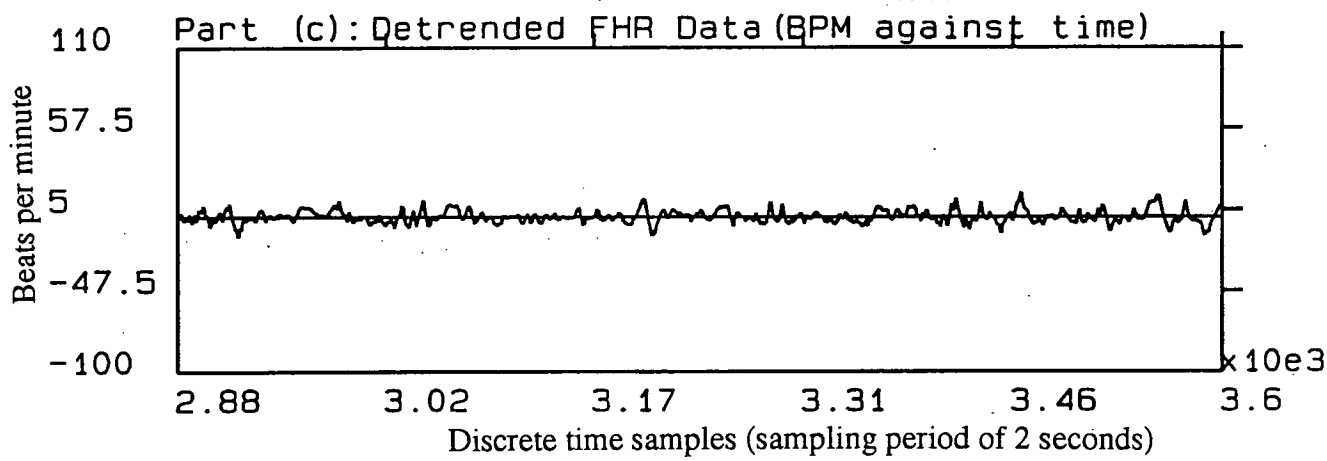
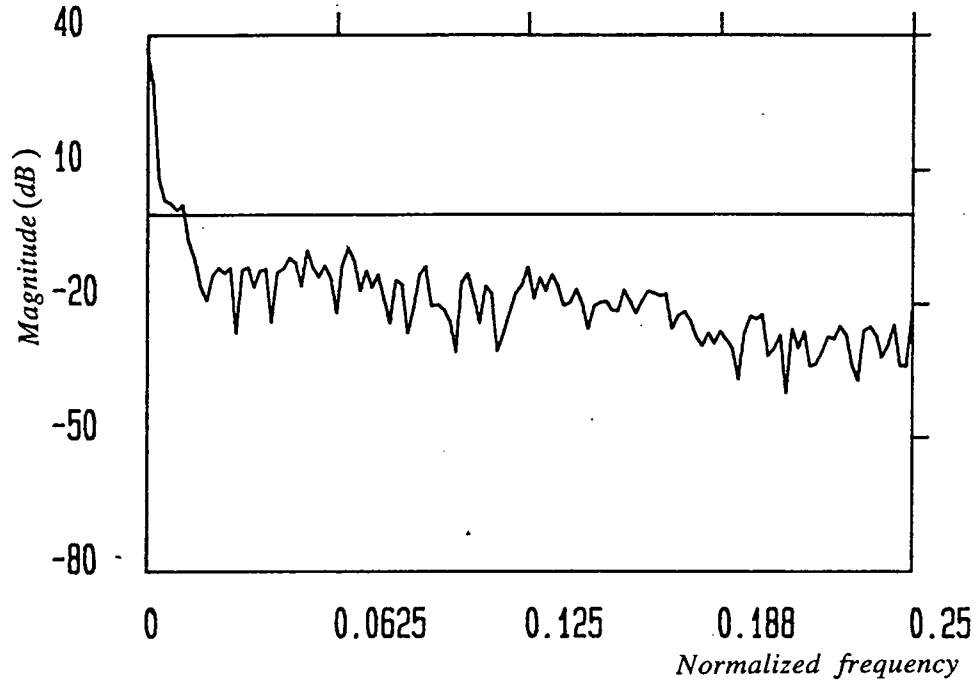


Figure 3.6: Part (a): Presents 24 minutes of averaged phonocardiographic FHR data, along with its estimated baseline.
Part (b): Shows the estimated baseline on its own.
Part (c): Shows detrended, mean stationary, averaged FHR data.

Part (d): FFT based power spectrum from the original FHR data.



Part (d): FFT based power spectrum from the detrended version.

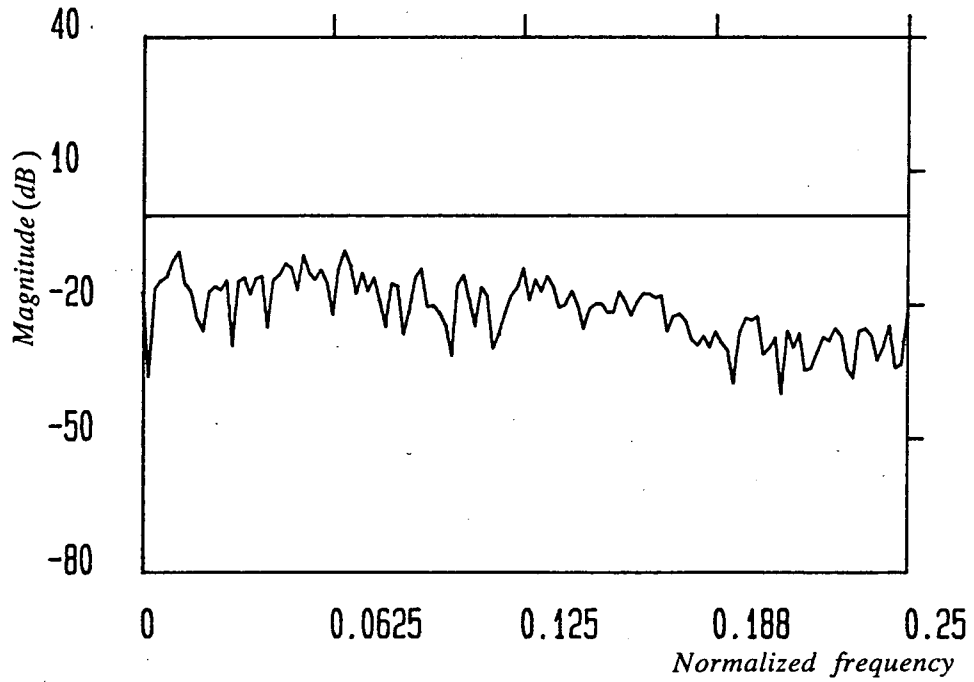


Figure 3.6: (d): FFT based power spectrum taken from the original 24 minutes FHR data segment (part (a)) and from its corresponding detrended version (part (c)), showing the effect of the trend removal operation on the spectral content of the FHR data (Magnitude (in dB) against normalized frequency).

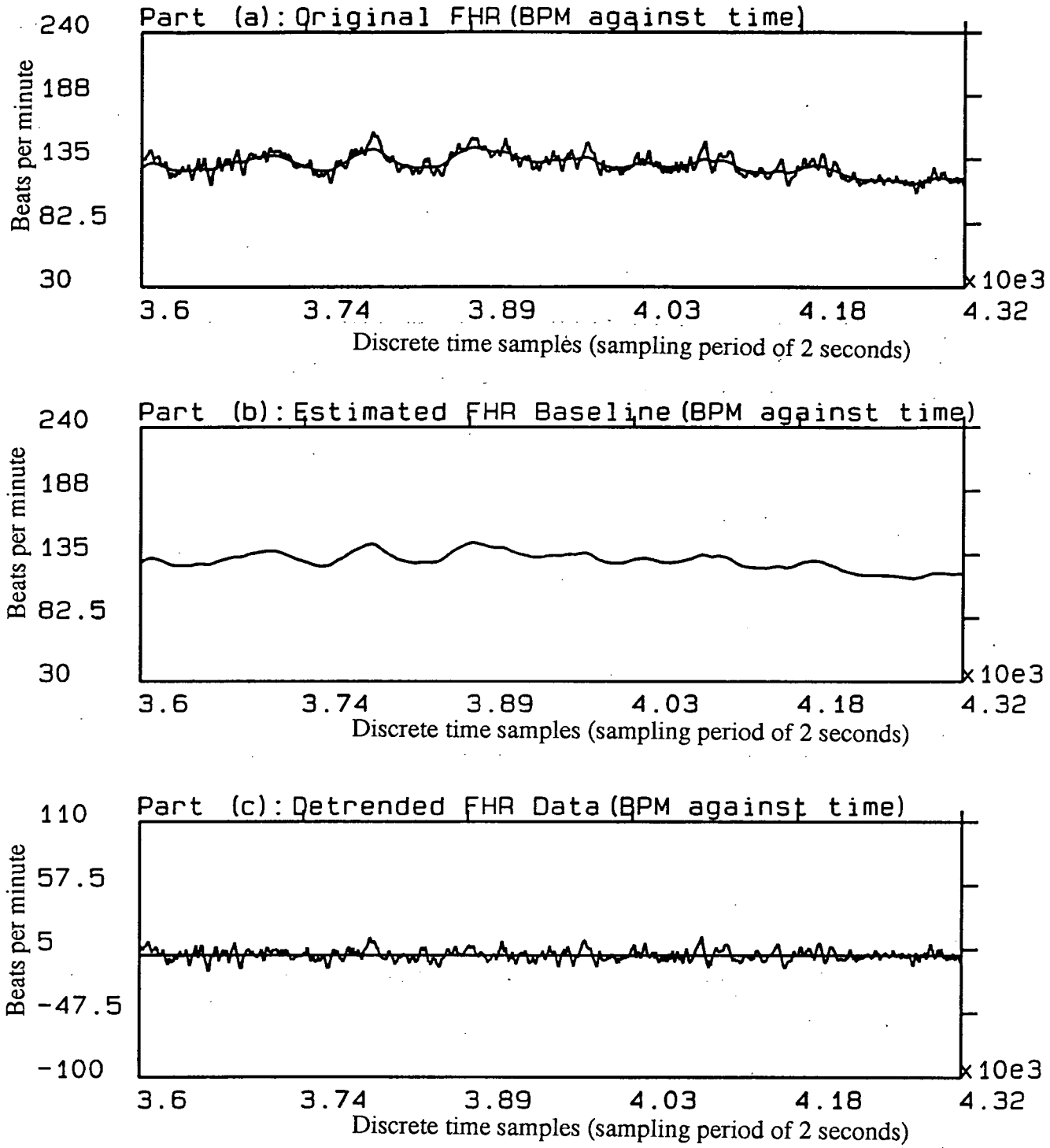
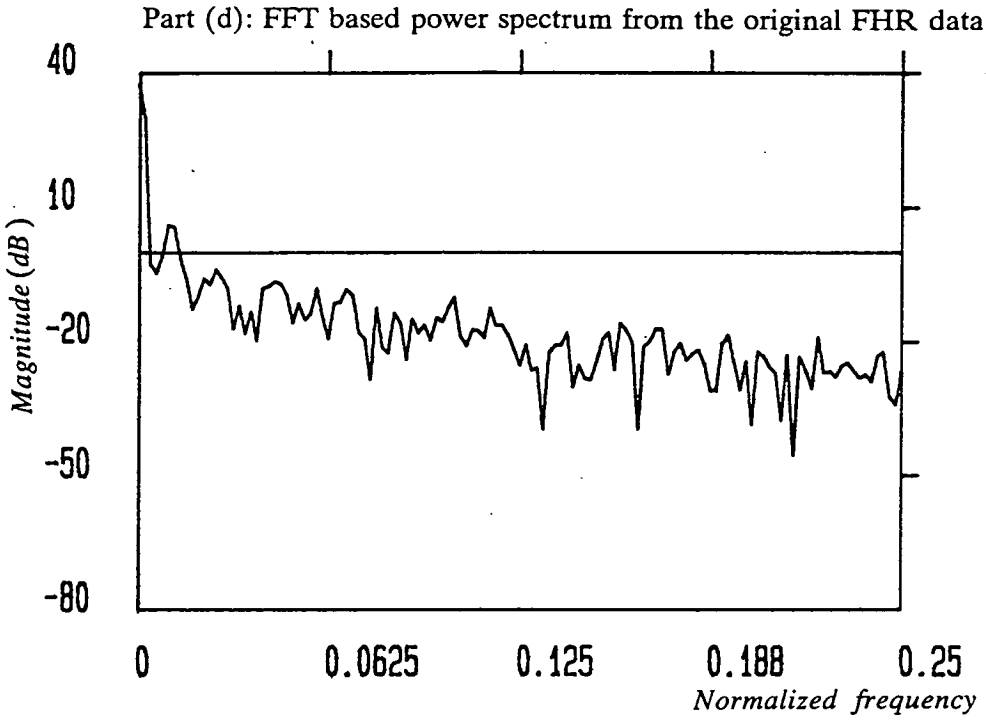


Figure 3.7: Part (a): Presents 24 minutes of averaged phonocardiographic FHR data, along with its estimated baseline.
Part (b): Shows the estimated baseline on its own.
Part (c): Shows detrended, mean stationary, averaged FHR data.



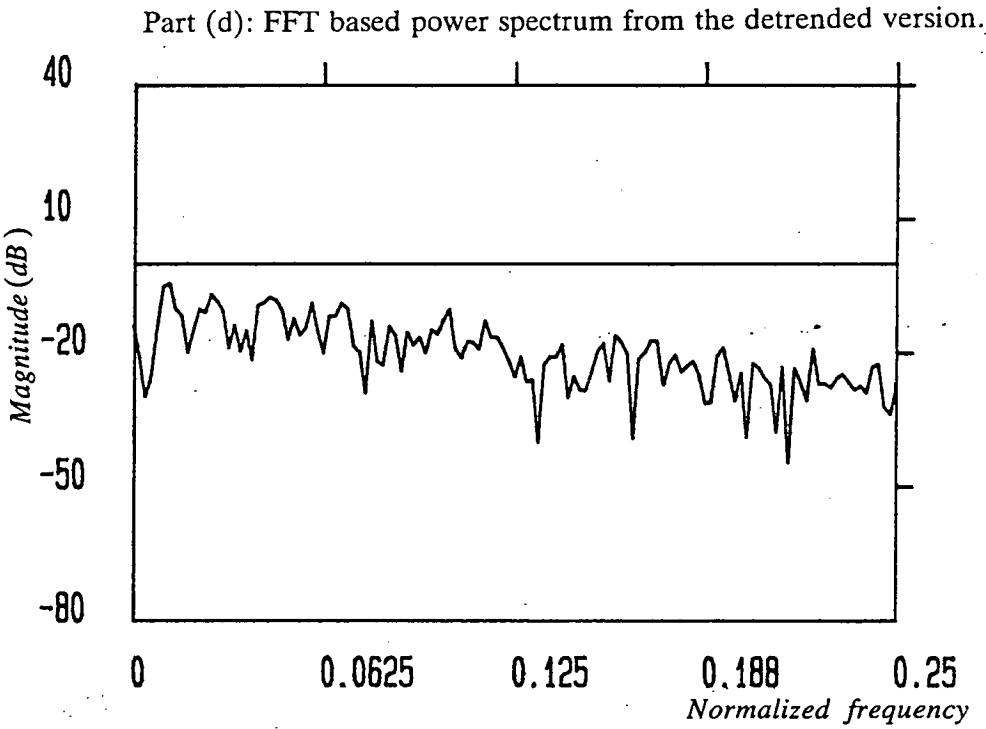


Figure 3.7: (d): FFT based power spectrum taken from the original 24 minutes FHR data segment (part (a)) and from its corresponding detrended version (part (c)), showing the effect of the trend removal operation on the spectral content of the FHR data (Magnitude (in dB) against normalized frequency).

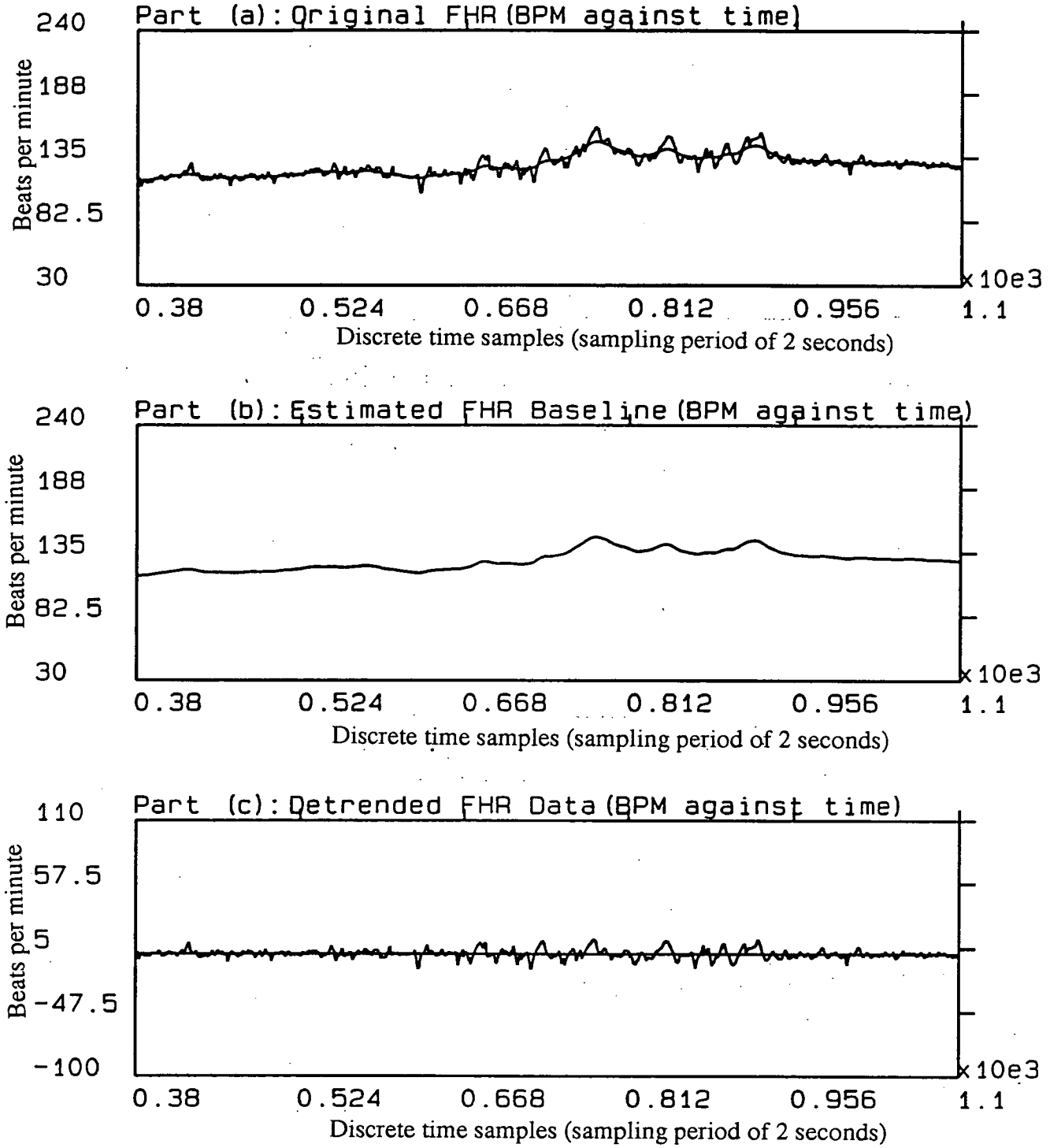


Figure 3.8: Part (a): Presents 24 minutes of averaged phonocardiographic FHR data, along with its estimated baseline.
Part (b): Shows the estimated baseline on its own.
Part (c): Shows detrended, mean stationary, averaged FHR data.

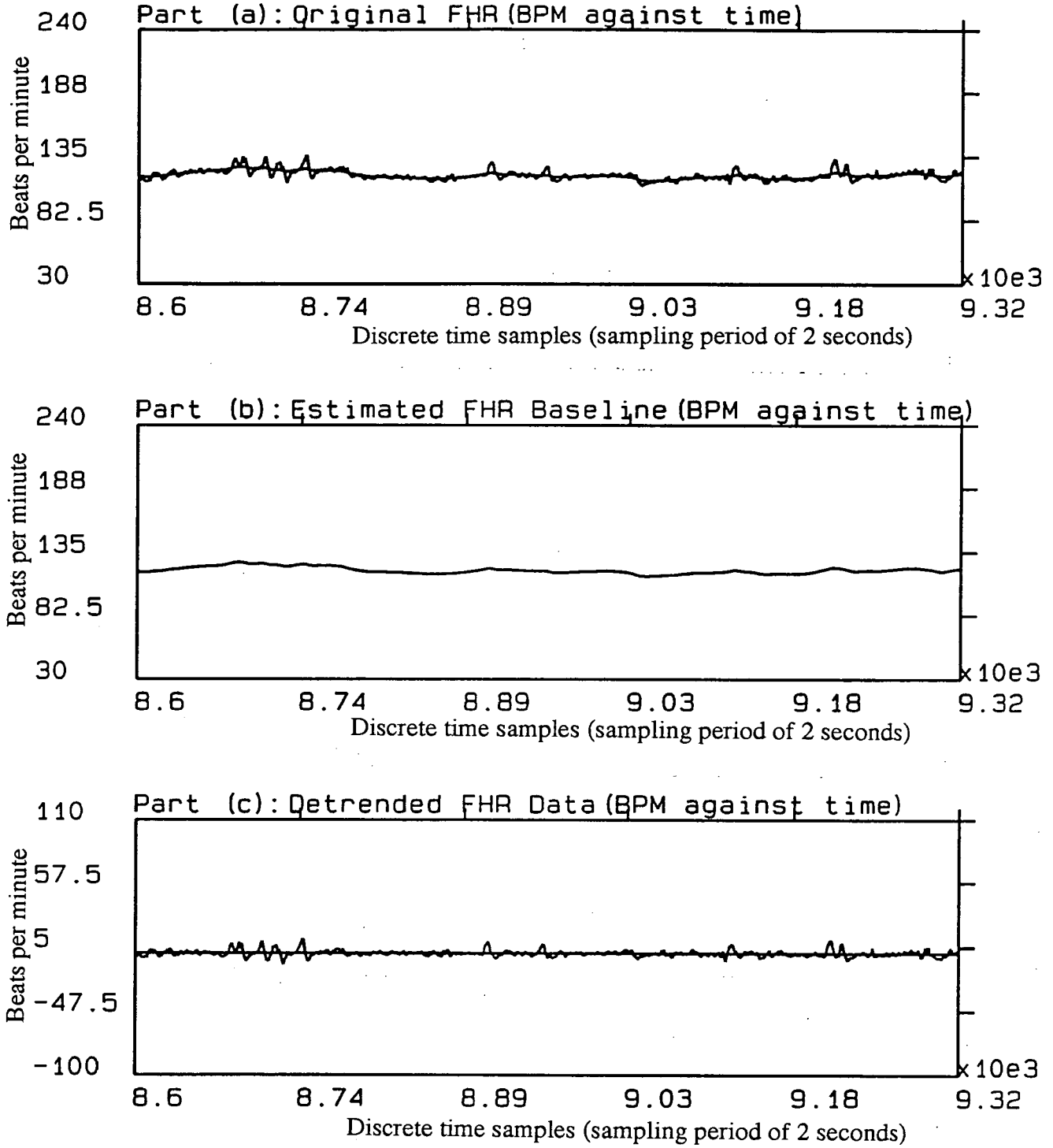
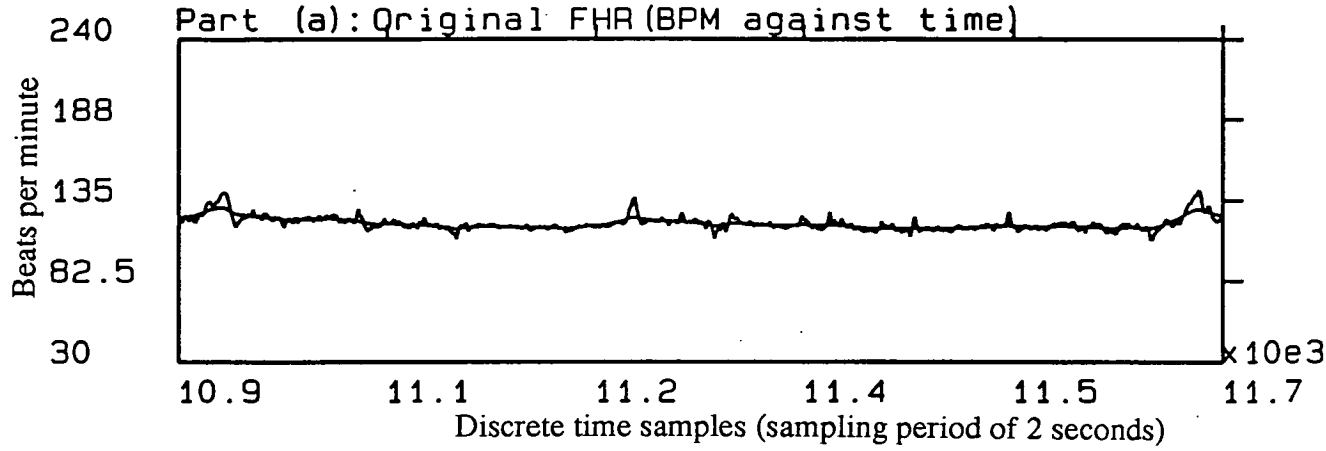
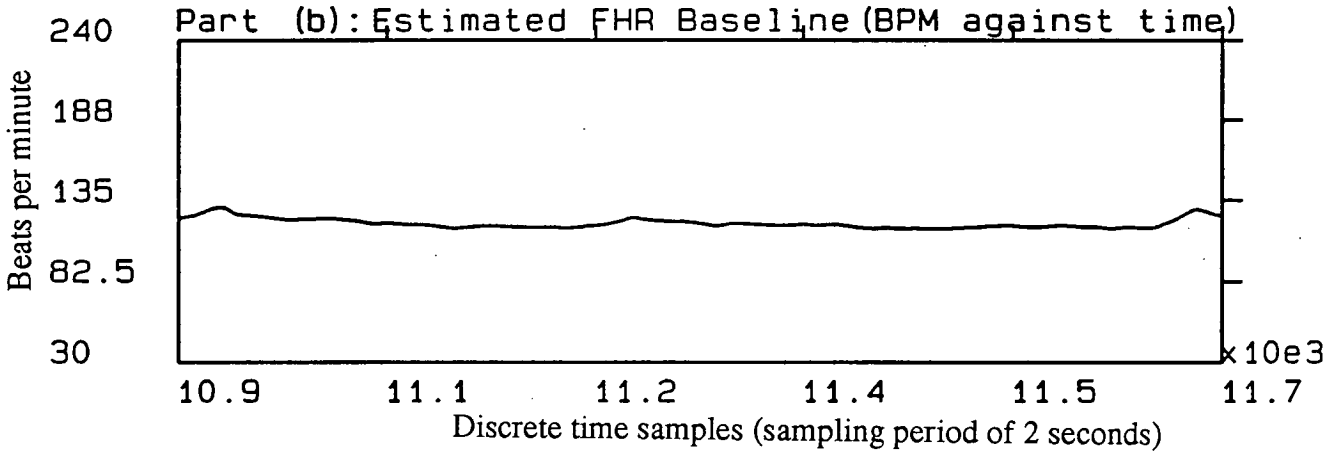


Figure 3.9: Part (a): Presents 24 minutes of averaged phonocardiographic FHR data, along with its estimated baseline.
Part (b): Shows the estimated baseline on its own.
Part (c): Shows detrended, mean stationary, averaged FHR data.





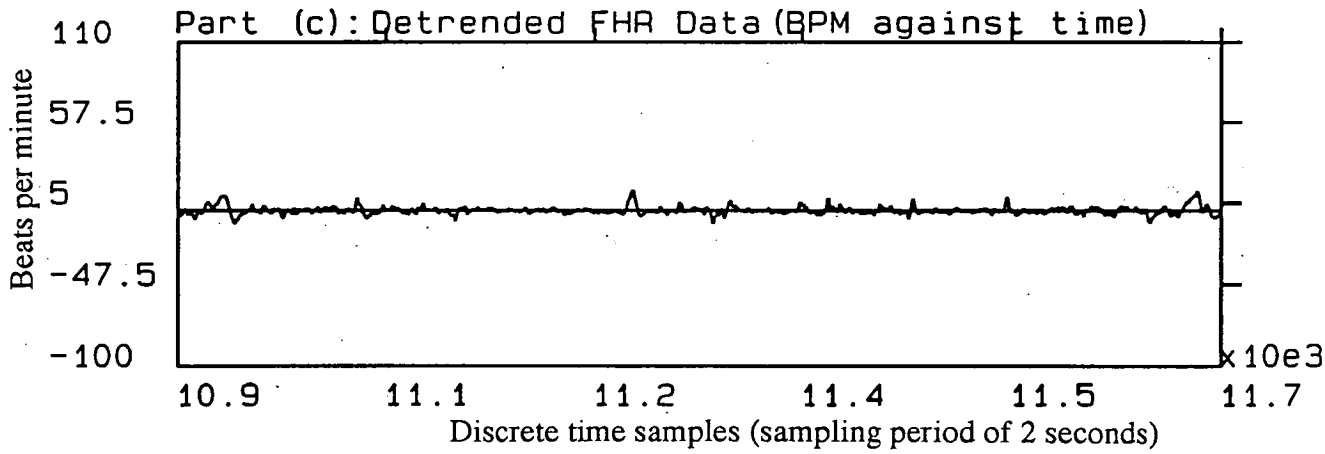


Figure 3.10: Part (a): Presents 24 minutes of averaged phonocardiographic FHR data, along with its estimated baseline.
Part (b): Shows the estimated baseline on its own.
Part (c): Shows detrended, mean stationary, averaged FHR data.

3.4 Statistical Comparison of the Original Series and its Detrended Version

As was mentioned in the previous chapter, analysis of variability through scalar type statistics (which all seem to be related to the standard deviation statistic), has not been very consistent or reliable in characterizing true FHR variability or FHR spread over the baseline.

The one-dimensional standard deviation statistic or related scalar statistical measures that have been applied in the past tend to base their computation on the original mean nonstationary FHR series, and this has been the main reason for their lack of success.

Trend has an enormous influence on the value of the standard deviation FHR variability. This renders the standard deviation statistic a useless quantity for representing true variability around the baseline, particularly when a pronounced first moment non-stationarity is present.

In the FHR recording shown in figure 3.5a, there was a series of long duration accelerations. The trend removal technique has included these long duration accelerations as part of the baseline. The variance of the original averaged FHR block (Fig. 3.5a) was 113.944 BPM^2, corresponding to a standard deviation of 10.674 BPM. However the variance of the corresponding detrended series (Fig. 3.5c) was only 17.507 BPM^2 corresponding to a standard deviation of 4.184 BPM.

Similarly the FHR recording shown in figure 3.6a, exhibited three long duration accelerations, which were included as part of the trend by the linear bi-directional filtering technique. Again the original FHR data had a high variance of 107.633 BPM^2 (standard deviation of 10.375 BPM), whereas its corresponding detrended version (Fig. 3.6c) had much lower variance of 17.182 BPM^2 (in standard deviation of

4.145 *BPM*). Similar large discrepancies existed in all other blocks of recordings analyzed, where they all exhibited a pronounced underlying trend or slow rhythm as seen in figures 3.5a to 3.8a.

In cases where there were no obvious strong underlying trends like the one shown in figure 3.9a and 3.10a, there was not much difference in the values of variance (or standard deviation) of the original and its corresponding detrended version (see table 3.1).

Table 3.1 presents the first and second moment statistical values of the original FHR time series, for the original and for the detrended versions shown in figures 3.5 to 3.10, making the above arguments clearer.

It can also be seen from table 3.1 that the estimated mean heart rate of the original (trended) time series is almost identical in value to the estimated mean of the corresponding trend or baseline time series.

3.5 Summary and Conclusion

The methodology which provided the indirect FHR time series, namely the phonocardiographic method was explained in detail. The nature of the FHR data that it provided was highlighted. The averaged FHR time series nature of the captured FHR data was stressed.

In this chapter the importance of baseline estimation and the need for the efficient removal of the baseline in a statistically unbiased manner as a pre-requisite for the objective numerical analysis of FHR time series has been made clear.

	Mean(BPM)		Variance (BPM^2)		Standard Deviation (BPM)	
Recording	Averaged FHR	Estimated Trend	Averaged FHR	Detrended FHR	Averaged FHR	Detrended FHR
Fig. 3.5	132.483	132.173	113.944	17.507	10.674	4.184
Fig. 3.6	126.442	126.228	107.633	17.181	10.375	4.145
Fig. 3.7	130.211	129.874	83.339	19.515	9.129	4.147
Fig. 3.8	128.838	128.680	91.366	12.614	9.558	3.552
Fig. 3.9	120.007	119.967	15.188	6.258	3.897	2.502
Fig. 3.10	119.989	119.892	21.612	6.432	4.649	2.536

Table 3.1: Mean, Variance and Standard Deviation of FHR for original (trended) and corresponding detrended versions. All six recordings were of 24 minutes duration.

A successful trend estimation and removal technique has been introduced, which can also be applied to any other biomedical time series. The technique used was based on a linear bi-directional autoregressive low-pass discrete digital filter.

The effect of mean nonstationarity (FHR trend) on the standard deviation statistic, which is the general scalar statistic used for variability analysis, has been investigated. As a result it was concluded that in order to represent the genuine spread of the variability over the baseline, through the use of the scalar standard deviation statistic (or indeed any related statistic), it is vitally important to use detrended FHR data instead of the original, raw data.

In the next chapter, detrended data will be used for variability analysis, using conventional scalar statistical analysis techniques, and the estimated trend with its corresponding rate of change along with the original data will be used for the detection of accelerations.

CHAPTER 4

STATISTICAL ANALYSIS OF FETAL HEART RATE VARIABILITY AND NUMERICAL DETECTION OF ACCELERATIONS

4.1 Introduction

In the previous chapter an important component of FHR, namely its baseline or trend was efficiently estimated. This trend must be removed prior to objective numerical analysis of the other two important diagnostic screening components of FHR, namely the FHR variability and accelerations. The detrimental effect of the FHR trend on the estimates of the values of the standard deviation statistic was investigated and demonstrated.

It was shown that, in order to analyze numerically the genuine variability, the slow underlying rhythm must be extracted. In other words, to properly compute the genuine fluctuations around the baseline or trend of the time series, detrended data must be used.

In this chapter a conventional scalar statistical analysis technique is applied to the detrended FHR data for the analysis of the variability component of the FHR. A new approach is presented, based on the use of a relative frequency histogram for displaying the general status of FHR variability in long term antepartum recordings.

Then a combination of different orders of moment based statistics are used to summarize effectively the histogram distribution shapes.

The other vital component of FHR which is used for determining the fetal medical status is the presence of accelerations. The presence of accelerations is always looked upon as a positive sign. However, there is little mention of successful numerical detection of accelerations in the published literature.

In this chapter a ruled based routine will be presented which makes use of the baseline information for detecting genuine accelerations. The technique scans two data sets simultaneously; the original FHR time series and its rate of change of the trend and is particularly useful when long term antepartum FHR recordings, need to be scanned for detection of accelerations.

4.2 Statistical Analysis of Variability

The detrended FHR series represents the true long term FHR variability, which is basically the main component of variability. As can be observed, the detrended series exhibits nonstationarity in the second moment (see part (c) of figures 3.5 to 3.10). Its variance continuously changes with time. In order to correctly analyze the time dependent behaviour of variability, the shortest possible observation epoch should be taken.

Statistically speaking, for an unbiased second order moment estimation, the number of samples in an observation window must be at least 30[49]. Any block of 30 or more samples is considered statistically as being a *large sampling* block. Ideally, however, the higher the number of samples in an observational segment, closer will be the statistical properties of the sample estimate to the "true population" statistics (i.e the estimator becomes more efficient).

Therefore the shortest observational epoch which could be used for standard deviation estimation was limited to one minute, exactly equivalent to a sampling block size of 30. This size of window was also short enough to have local statistical stationarity in the second moment.

Effectively one can look upon the one minute block by block standard deviation estimate of detrended data, as a one-dimensional way of data reduction, being able to represent quantitatively progressive changes in the long term variability.

The question that arises is, what is the best way of analyzing and representing these scalar statistical estimates. The key to the answer lies in the way in which the level of variation in long term variability has been grouped and used by the medics[50]. Considering the latter, a relative frequency histogram presentation has been adopted for display of the one minute epoch standard deviation statistics estimated from the detrended FHR data [49, 51].

4.3 The Classification of Variability

The general requirement for normal variability is a peak-to-peak fluctuation of more than 5 BPM. This is equivalent to having a time domain standard deviation of more than 3 milliseconds or a frequency domain standard deviation of more than 0.83 BPM.

Once the trends have been removed, the variability may be classified by subdivision into five groups[1, 10, 50, 52, 53], these are:

1. Silent pattern (Peak-to-peak fluctuations of less than 2 BPM, corresponding to a standard deviation of less than 1 ms in the time domain or 0.28 BPM in the frequency domain).

2. Minimal oscillations (Peak to peak fluctuations of 3 to 5 BPM, corresponding to a 1 to 3 ms standard deviation in the time domain or 0.28 to 0.83 BPM in the frequency domain).

3. Small oscillations (Peak to peak fluctuations of 5 to 10 BPM, corresponding to a 3 to 7 ms standard deviation in the time domain or 0.83 to 1.94 BPM in the frequency domain).

4. Moderate oscillations (Peak to peak fluctuations of 10 to 25 BPM, corresponding to a 7 to 10 ms standard deviation in the time domain or 1.94 to 2.77 BPM in the frequency domain).

5. Large oscillations (Peak to peak fluctuations of more than 25 BPM, corresponding to a standard deviation of more than 10 ms in the time domain or 2.77 BPM in the frequency domain).

In the following example a recording of phonocardiographic derived FHR of approximately 8 hours duration has been detrended and converted into time domain (ms) in order to use the FHR variability analysis groupings explained above (see figure 4.1).

The recording has been segmented into 1 minute intervals, and their corresponding standard deviation estimates taken. Altogether 472 standard deviation statistical values were derived from this recording.

4.4 Histogram Display of The Computed Standard Deviation Values

To obtain an accurate representation and a good picture of the distribution of variability in a long term antepartum recording, in terms of the FHR Variability groupings introduced above, a relative frequency histogram of the one minute epoch stan-

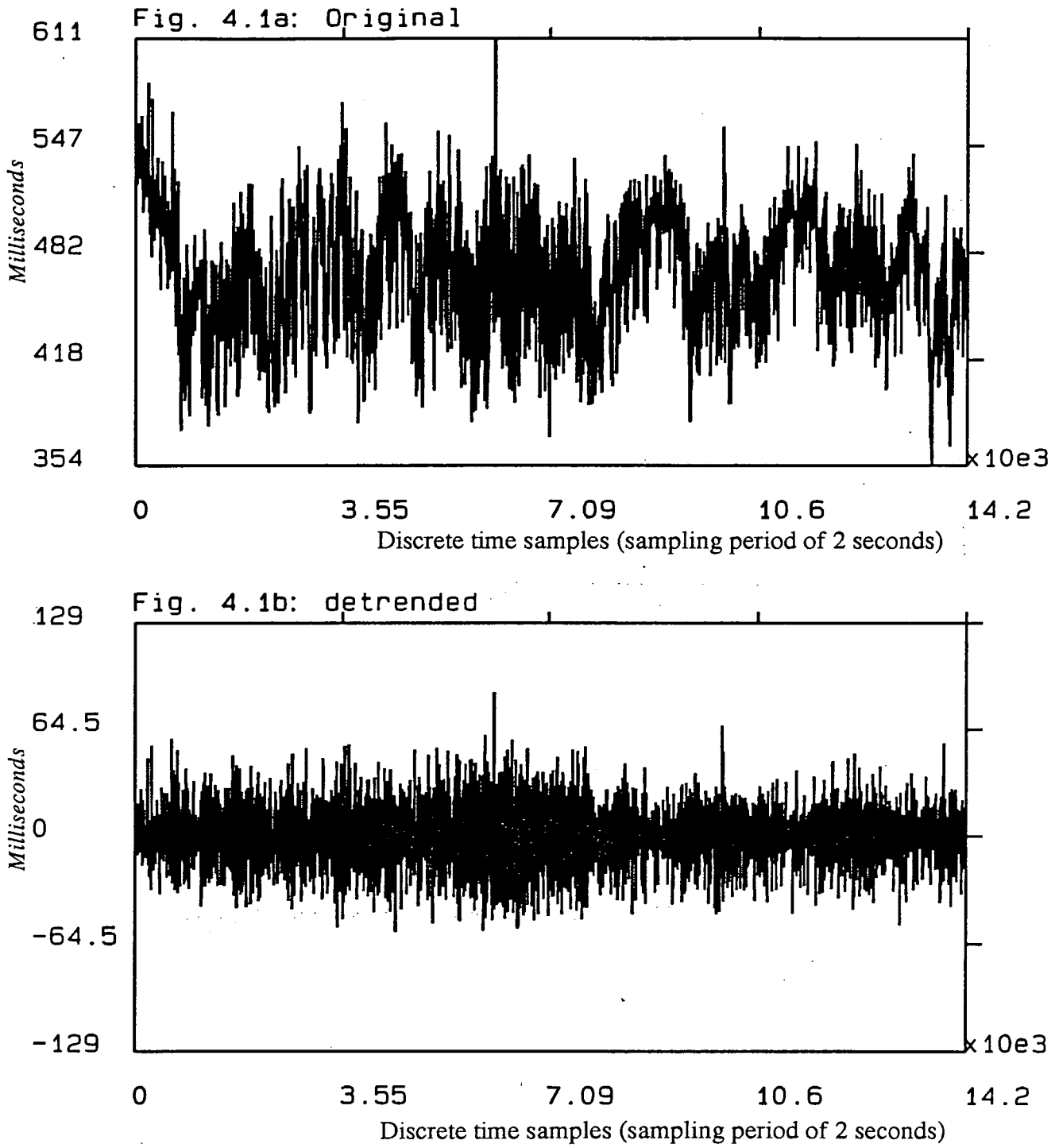


Figure 4.1: Part (a): Showing an eight hour long averaged Fetal Heart Period phono-cardiographic recording. The FHP values shown in this plot are all in *ms* (time domain periodic rate) against time.
Part (b): Shows the detrended version of the eight hour long recording shown in part (a), again the FHP values are presented in *ms*

dard deviation values computed was configured.

The maximum standard deviation estimate that has been observed in this data was 34 ms, so the range chosen for the histogram analysis was from 0 to 36 ms. This range was then subdivided into 36 equally spaced intervals of 1 ms duration.

The number of occurrences of standard deviation values falling in any of these bins was then scanned for, and finally the relative frequency values for the individual bins was computed[49, 51].

Figure 4.2a and 4.2b show the relative frequency histograms of one minute block standard deviation statistics derived from the original FHR time series and from its corresponding detrended version respectively. The effect of detrending is very much apparent in these two histogram plots. In the case of figure 4.2(b) (detrended data), a much bigger percentage of estimated standard deviation values occupy the lower bin intervals than in figure 4.2(a) (raw data).

The standard deviations computed from one minute epochs of the detrended data constitutes a good estimate of the variability over the baseline over that short interval of time. When presented using a relative frequency histogram these statistics should constitute, a useful tool for objective numerical analysis of FHR variability in long term antepartum FHR recordings.

Some medical conditions may have characteristic histogram parameters which would permit early detection of the condition. The histogram parameters should be regarded as constituting a vector. Then clustering techniques based on statistical measurements such as Mahalanobis distances can be applied[54]. This is suggested as further work.

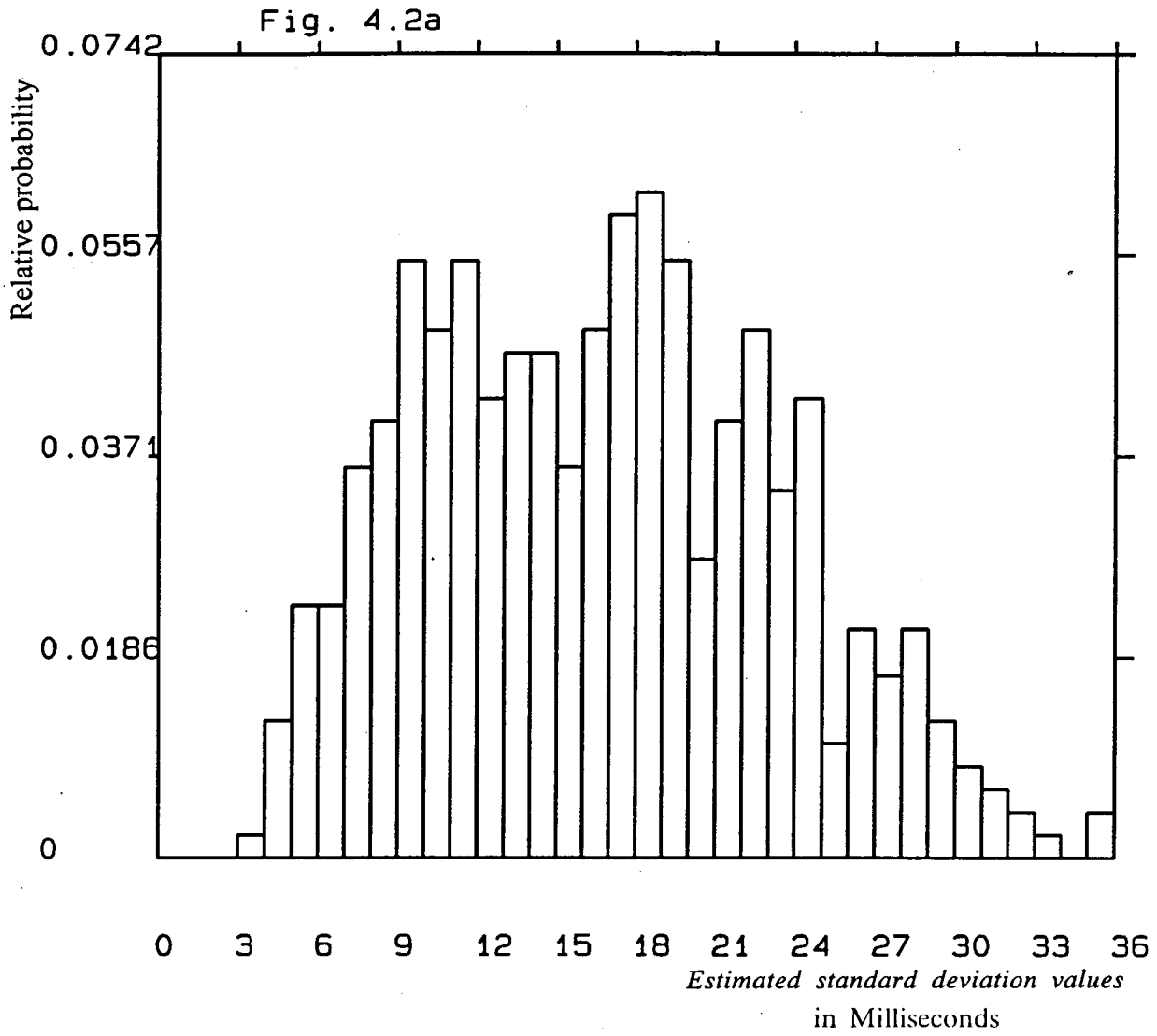


Figure 4.2: Part (a): Relative frequency histogram of the standard deviations computed from one minute blocks of the original FHP time series shown in part (a) of figure 4.1. The shape of the histogram is summarized via the following moment based scalar statistics:

$\bar{x} = 16.7804$, *where its standard error* $S_{\bar{x}} = s/\sqrt{N} = 0.3211$,

$s = 6.9771$, *where its standard error* $S_s = s/\sqrt{2N} = 0.2271$,

$\varsigma = 0.4521$, *where its standard error* $S_\varsigma = \sqrt{6/N} = 0.1127$,

$\kappa = -0.0277$, *where its standard error* $S_\kappa = \sqrt{24/N} = 0.2255$.

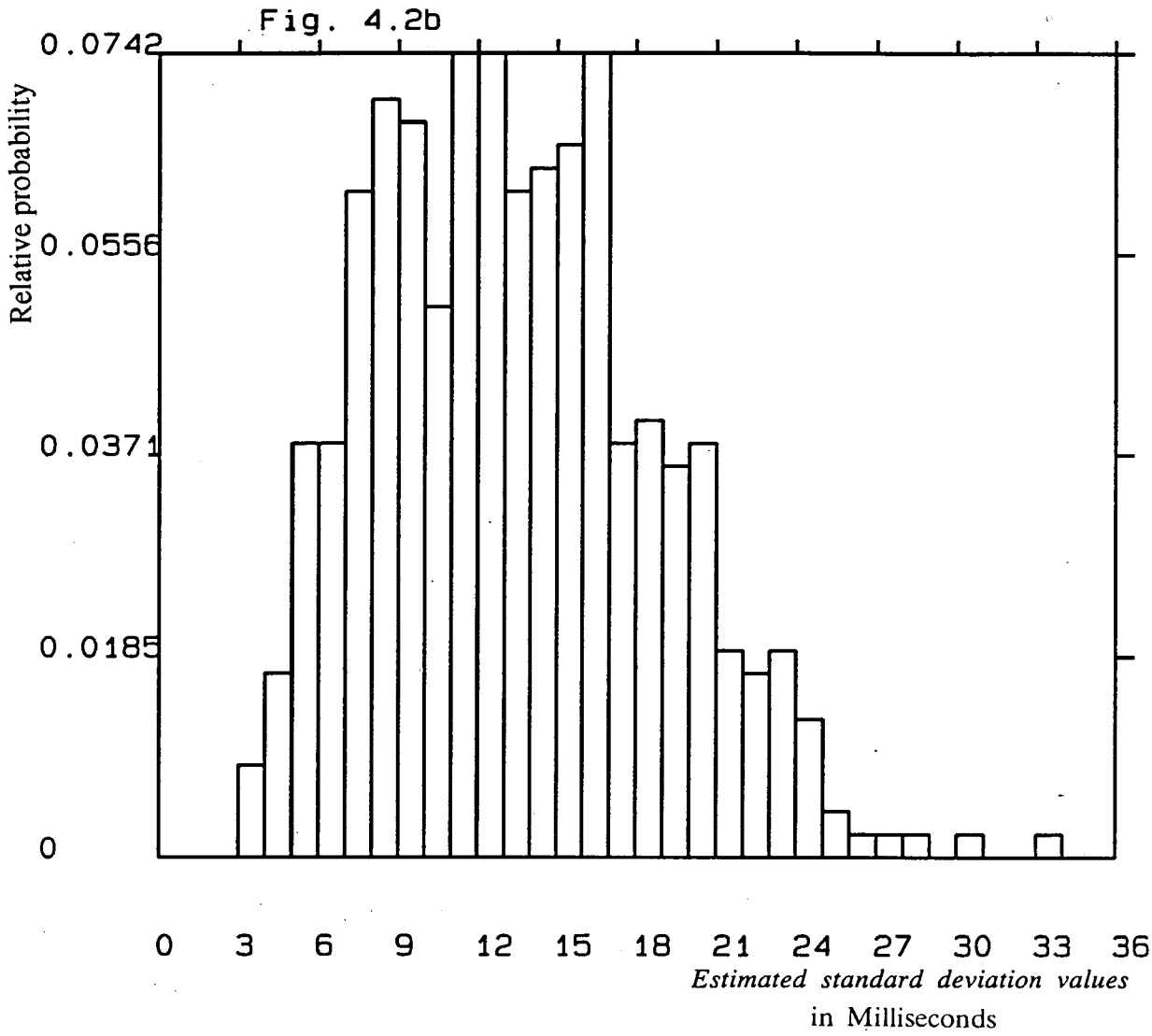


Figure 4.2: Part (b): Relative frequency histogram of the standard deviations computed from one minute blocks of the detrended FHP time series shown in part (b) of figure 4.1. The shape of the histogram is summarized via the following moment based scalar statistics:

$\bar{x} = 13.4624$, *where its standard error* $S_{\bar{x}} = s/\sqrt{N} = 0.2413$,

$s = 5.2413$, *where its standard error* $S_s = s/\sqrt{2N} = 0.1706$,

$\varsigma = 0.4426$, *where its standard error* $S_\varsigma = \sqrt{6/N} = 0.1127$,

$\kappa = -0.1377$, *where its standard error* $S_\kappa = \sqrt{24/N} = 0.2255$.

4.4.1 Numerical Representation of Histogram Shapes

It would be useful if the histogram shapes derived previously for the variability analysis could be represented by a set of numerical values. This data compression will be of particular value when one is dealing with large numbers of derived distribution and for classification purposes as discussed in the above section. The histogram shown in figure 4.2a, may be modeled as an asymmetric unimodal distribution. The peak of the distribution is biased towards the left hand side, and its upper tail is longer than the lower tail. This is referred to as being skewed to the right, or positively skewed.

There are two statistical quantities based respectively on the third and fourth moments of the sample estimates, which can numerically represent the general shape of relative frequency histograms. These are referred to as the coefficient of skewness (ς) and coefficient of kurtosis (κ) respectively [49, 51, 55, 56].

Both of these statistics, will be equal to zero if the distribution exhibited by the relative frequency histogram is Gaussian. It should also be added that these higher moment based (above second) statistics can best be used if and only if the sample size is very large (larger than 150 samples for ς and larger than 1000 samples for κ)[57, 58].

Skewness is the degree of asymmetry, or departure from symmetry of a distribution. If the frequency curve (the histogram) of a distribution has a longer tail to the right of the central maximum than to the left, then the distribution is said to be skewed to the right or to have positive skewness. If the reverse is true, it is said to be skewed to the left or to have negative skewness. In the case of the histograms shown in figures 4.2a and 4.2b, the distribution can be said to have positive skewness.

There are a number of ways of measuring skewness, but the simplest and most important way is to use the third moment about the mean of the sample estimate and to express it in dimensionless form as is given by:-

$$\varsigma = \frac{\sum_{k=1}^{N}(x_k - \bar{x})^3/N}{s^3}, \quad \textit{with a standard error} \quad S_\varsigma = \sqrt{\frac{6}{N}}, \tag{4.1}$$

where x_k denotes the k-th of the N values assumed by a variable x, s denotes the sample standard deviation, and $\bar{x}$ denotes the sample mean, calculated from the N values that are assumed by x. In this application, x_k represents the k-th of the N standard deviation values estimated from one minute segments of detrended FHR time series.

Kurtosis is the degree of peakedness of a distribution, usually taken relative to a normal distribution. A distribution having a relatively high peak is called *leptokurtic*, while the one which is flat-topped is called *platykurtic*. The normal distribution (which is not very peaked or very flat-topped) is called *mesokurtic*.

To measure the kurtosis, the fourth moment about the mean is used and is given by:

$$\mathrm{K} = \frac{\sum_{k=1}^{N}(x_k - \bar{x})^4/N}{s^4} \tag{4.2}$$

For a normal distribution, $\mathrm{K} = 3$. For this reason the kurtosis is usually defined by:-

$$\kappa = \mathrm{K} - 3, \quad \textit{with a standard error} \quad S_\kappa = \sqrt{\frac{24}{N}}, \tag{4.3}$$

which is positive for a leptokurtic distribution, negative for a platykurtic distribution, and zero for the normal (mesokurtic) distribution. In figure 4.2b, on visual inspection, the distribution exhibits positive skewness and is neither peaked nor flat-topped, and this is reflected in the computed positive value for ς and close to zero value for κ shown below figure 4.2b.

Although it must be said, that less significance can be placed on the estimated value for κ shown in figures 4a and 4b, because of its relatively large standard error. This is especially critical in the examples shown, where the estimated value for κ were close to zero. The only way to increase the significance of these (close to zero) estimates of κ, would be to increase, as far as possible, the sample size used in the estimate. This results in smaller standard error and therefore a more efficient statistical estimate. However, when the distribution of the available data is pronouncedly peaked or pronouncedly flat, even for a relatively small sample size (as in the case of the examples shown in figure 4.2 (N=472)) the computed κ value will be large enough in comparison to its standard error to support the hypothesis of a peaked or flat distribution respectively.

In summary using ς and κ together with the sample mean ($\bar{x}$) and standard deviation (s) statistics, the relative frequency histogram can be represented numerically. In this way the previously described FHR variability information can be numerically represented by just four moment based statistical values.

This compact numerical representation is of value when hundreds of such data sets need to be analyzed and classified.

4.5 Numerical Detection of Accelerations

As mentioned in the previous chapter, in the nonstress monitoring of patients, detection of accelerations in the FHR-trace is seen as a positive indicator of fetal health.

Obstetricians interpret as a sign of normality the detection of five or more genuine accelerations per hour from FHR recordings made during the antepartum period[21]. These accelerations are associated with spontaneous fetal movements.

In current routine clinical practice accelerations are looked for visually during the antepartum monitoring of FHR. This tends to be an unreliable and inaccurate way of analyzing accelerations, especially when one is dealing with long durations of recordings. Much more consistent results could be obtained by computer analysis.

One certainly would not be able to pin point visually the exact start, peak and end of an acceleration with their corresponding values. So the numerical detection and quantification of FHR accelerations was an important issue which had to be addressed.

The objective of this part of the work was rapid, accurate and consistent detection and quantification of accelerations in long FHR traces by the application of appropriate numerical techniques.

With the use of the estimated baseline discussed previously and the application of a rule based routine, this objective has been achieved.

The routine bases its detection of accelerations on the general rule that an acceleration is *genuine* if and only if there has been a rise of FHR of at least 15 BPM above the *steady state FHR baseline* (see figure 3.2), and maintenance of this rise for longer than 15 seconds, before falling back to a lower point where the baseline gradient will be zero[3, 22, 36].

In fact this is the general rule used to scan one hour of data for accelerations during an antepartum non-stress test (NST). The FHR is termed *reactive* if five or more of these accelerations are detected in one hour. Following such a strict rule, makes the task of visually detecting accelerations even more difficult.

4.5.1 Methodology Behind the Detection of Accelerations

The rule-based routine devised makes use of the baseline information, its corresponding computed rate of change and the original recorded averaged FHR time series to achieve its objective.

The rate of change of the baseline is approximated by the following first order finite difference equation:-

$$ROAD[k] = \frac{BAS[k+1]-BAS[k]}{T} \qquad (4.4)$$

where ROAD[k] represents the rate of accelerations (rises) away from and decelerations (falls) towards the steady state FHR baseline at the k-th sample, and BAS[k] the estimated baseline (or trend) at k-th sample and T the sampling period.

Figures 4.3 to 4.8 are examples showing the original FHR with superimposed estimated baseline and the rate of change of the baseline (estimated trend). It is clearly observable that for each FHR acceleration there is a distinct common pattern.

Firstly for every FHR rise from the steady state FHR baseline (where the approximate rate of change of the baseline is zero) there is always an FHR fall towards the steady state FHR baseline.

Also, where the FHR rise related to an FHR acceleration reaches a steady state peak value, the rate of change of the baseline during this period will also have an approximate zero value as well. These rules together with the rule of what constitutes a *genuine* FHR acceleration generally make up the structure of the devised routine which detects FHR accelerations.

For each FHR acceleration, in figures 4.3 to 4.8, the corresponding rate of change pattern starts from a value of zero; passes a heuristically found positive threshold

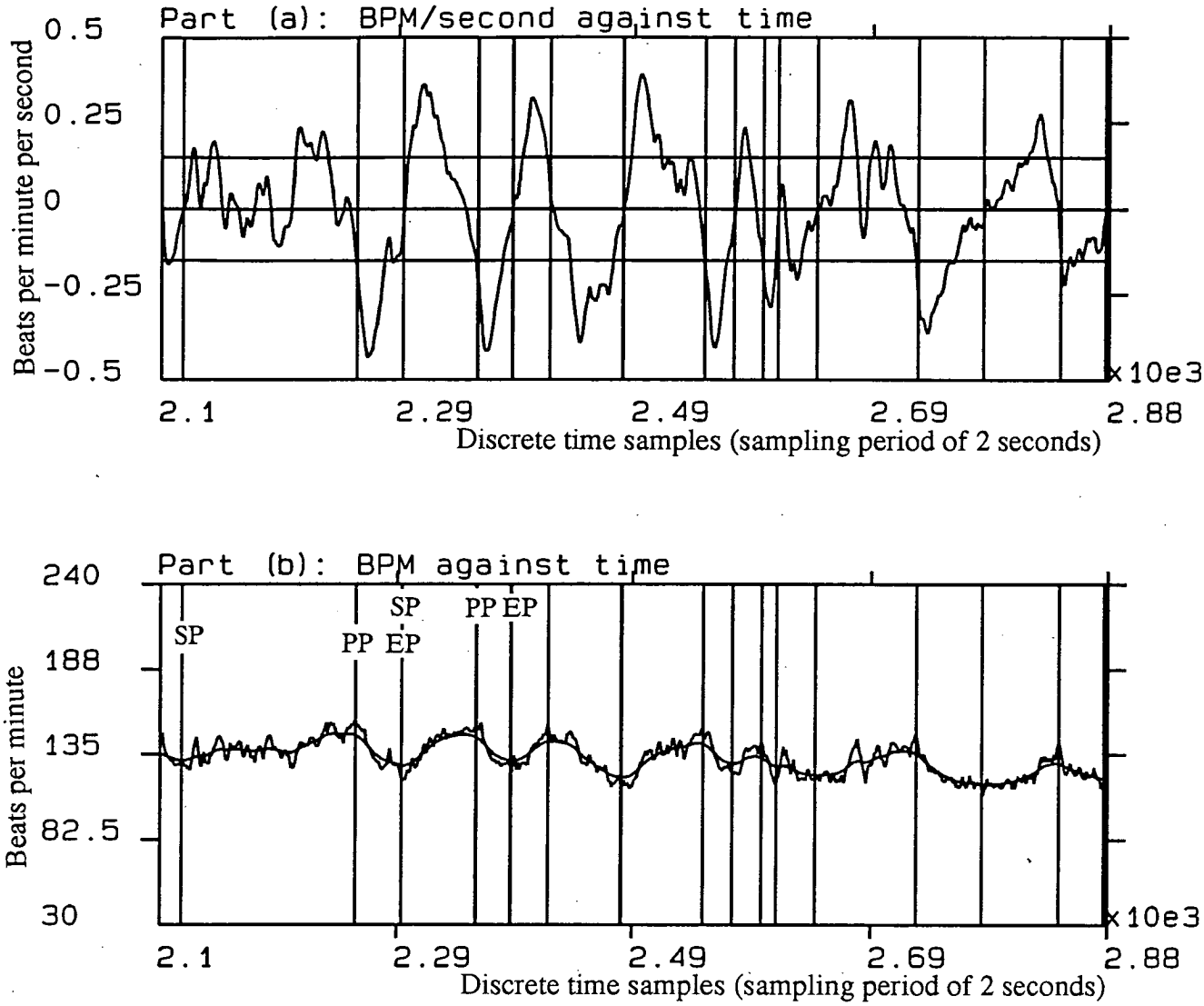


Figure 4.3: Part (a): Represents the rate of change of the baseline, shown in part (b) of the averaged FHR. The corresponding vertical lines drawn in (a) and (b) mark the points associated to starting (SP), peak (PP) and ending (EP) point of accelerations detected by the rule based routine. For example, in the case of this FHR data window shown, the routine will provide the following numerical information on each detected acceleration:

SPv: FHR value at SP; PPv: FHR value at PP; EPv: FHR value at EP; Length: duration of accelerations presented in terms of number of samples.
All the SP, PP, and EP values quoted here are presented in terms of sampling point position in the FHR data File being analyzed.

No.	SP.	SPv.(BPM)	PP.	PPv.(BPM)	EP.	EPv.(BPM)	Length.
1	2118	129.702	2261	155.394	2299	119.606	181
2	2299	119.606	2361	152.925	2389	125.875	90
3	2389	125.875	2420	153.230	2481	116.354	92
4	2481	116.354	2548	150.853	2572	124.265	91
5	2572	124.265	2596	142.845	2608	116.566	36
6	2640	120.474	2722	147.227	2776	113.612	136
7	2776	113.612	2840	141.448	2877	115.649	101

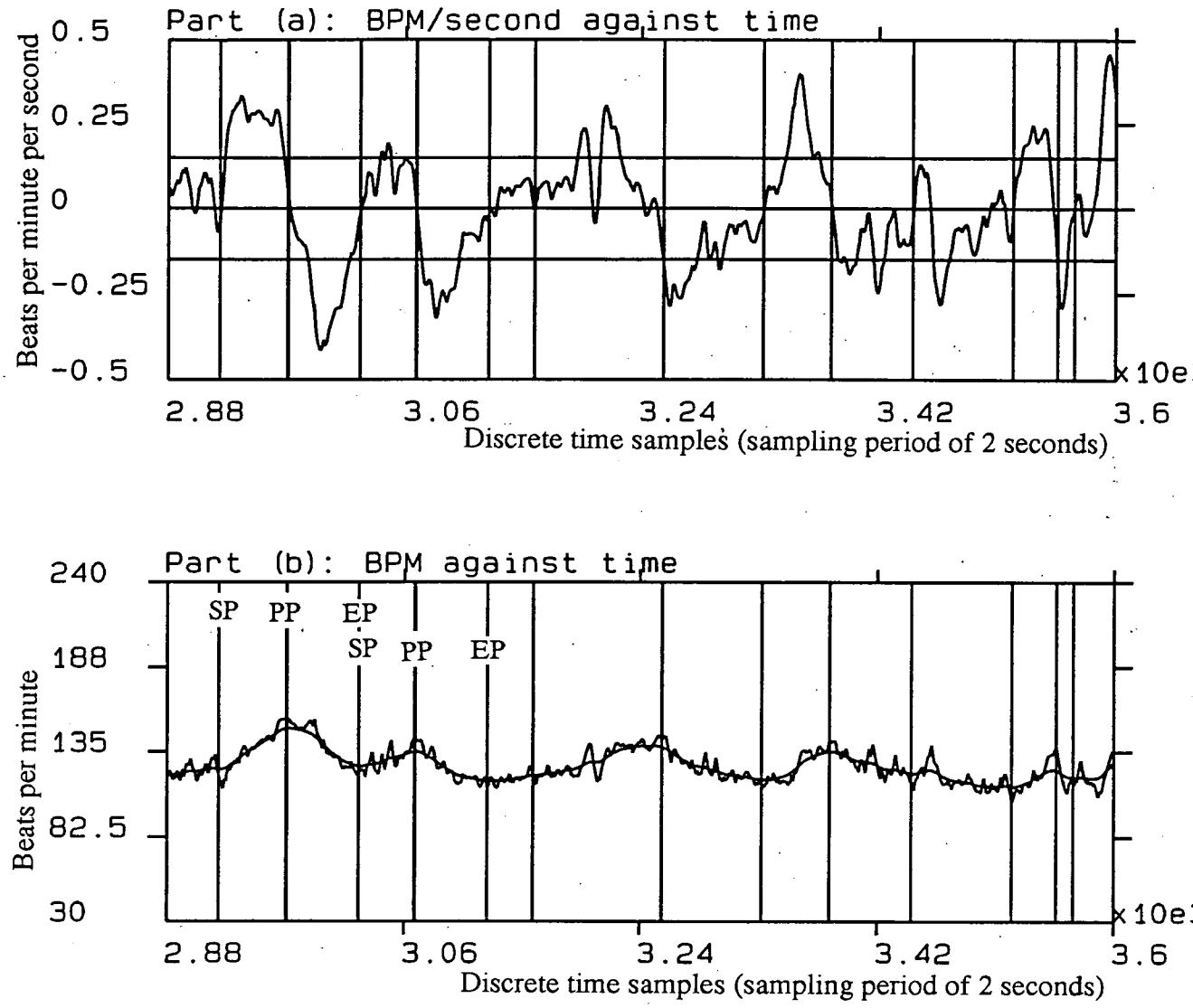


Figure 4.4: Part (a): Represents the rate of change of the baseline, shown in part (b) of the averaged FHR. The corresponding vertical lines drawn in (a) and (b) mark the points associated to starting (SP), peak (PP) and ending (EP) point of accelerations detected by the rule based routine. For example, in the case of this FHR data window shown, the routine will provide the following numerical information on each detected acceleration:

SPv: FHR value at SP; PPv: FHR value at PP; EPv: FHR value at EP; Length: duration of accelerations presented in terms of number of samples.
All the SP, PP, and EP values quoted here are presented in terms of sampling point position in the FHR data File being analyzed.

No.	SP.	SPv.(BPM)	PP.	PPv.(BPM)	EP.	EPv.(BPM)	Length.
1	2919	120.052	2970	155.486	3025	123.596	106
2	3025	123.596	3068	142.809	3124	112.843	99
3	3159	120.133	3257	145.368	3332	114.605	173
4	3332	114.605	3384	142.190	3446	116.573	114
5	3522	104.720	3556	137.691	3569	116.007	47

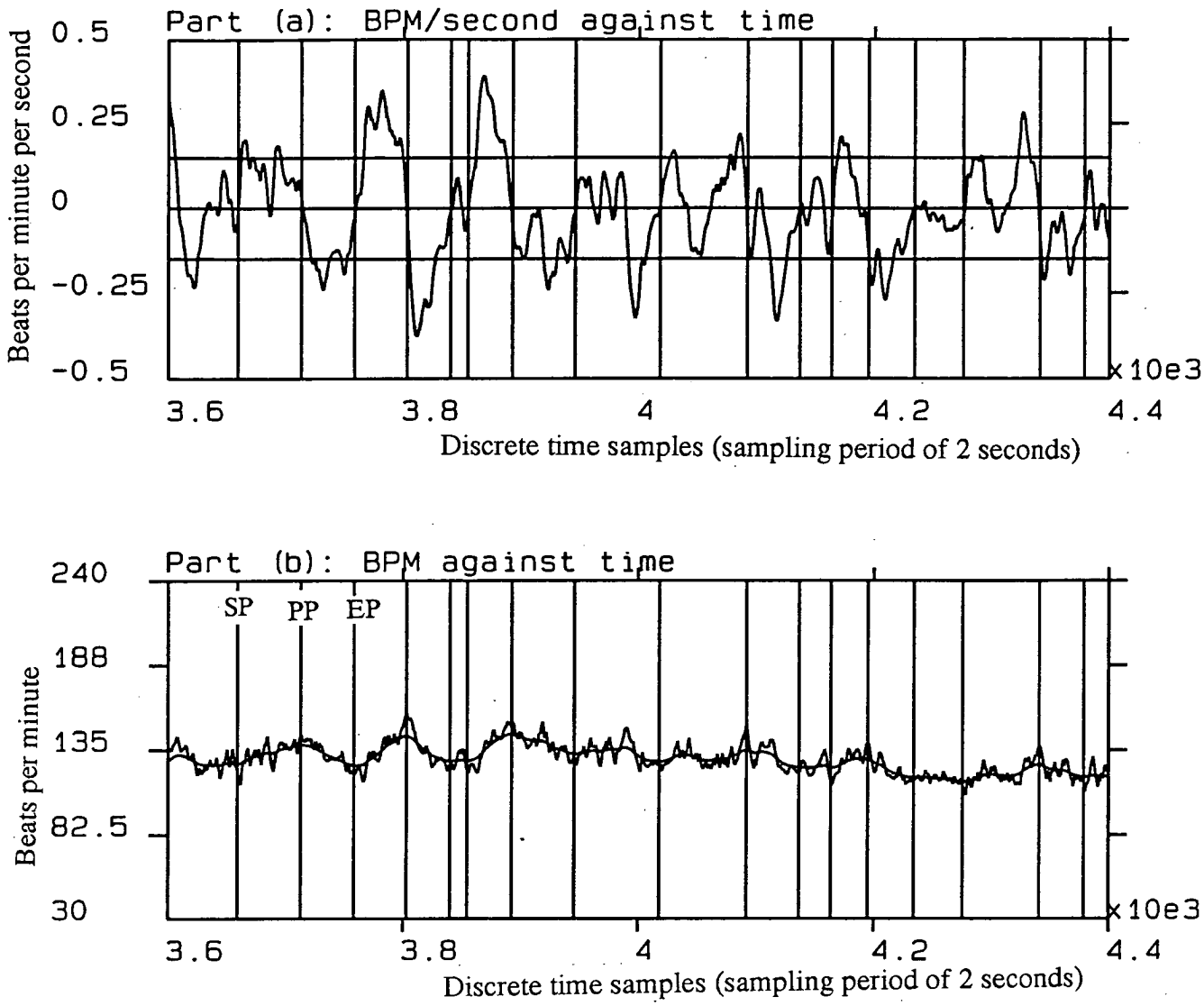


Figure 4.5: Part (a): Represents the rate of change of the baseline, shown in part (b) of the averaged FHR. The corresponding vertical lines drawn in (a) and (b) mark the points associated to starting (SP), peak (PP) and ending (EP) point of accelerations detected by the rule based routine. For example, in the case of this FHR data window shown, the routine will provide the following numerical information on each detected acceleration:

SPv: FHR value at SP; PPv: FHR value at PP; EPv: FHR value at EP; Length: duration of accelerations presented in terms of number of samples.
All the SP, PP, and EP values quoted here are presented in terms of sampling point position in the FHR data File being analyzed.

No.	SP.	SPv.(BPM)	PP.	PPv.(BPM)	EP.	EPv.(BPM)	Length.
1	3659	120.602	3712	144.423	3758	123.724	99
2	3758	123.724	3803	158.080	3840	126.146	82
3	3855	123.536	3893	152.800	3946	131.462	91
4	4018	125.777	4093	150.073	4137	121.611	119
5	4164	115.674	4195	139.453	4234	117.197	70
6	4275	112.790	4341	137.479	4379	113.444	104

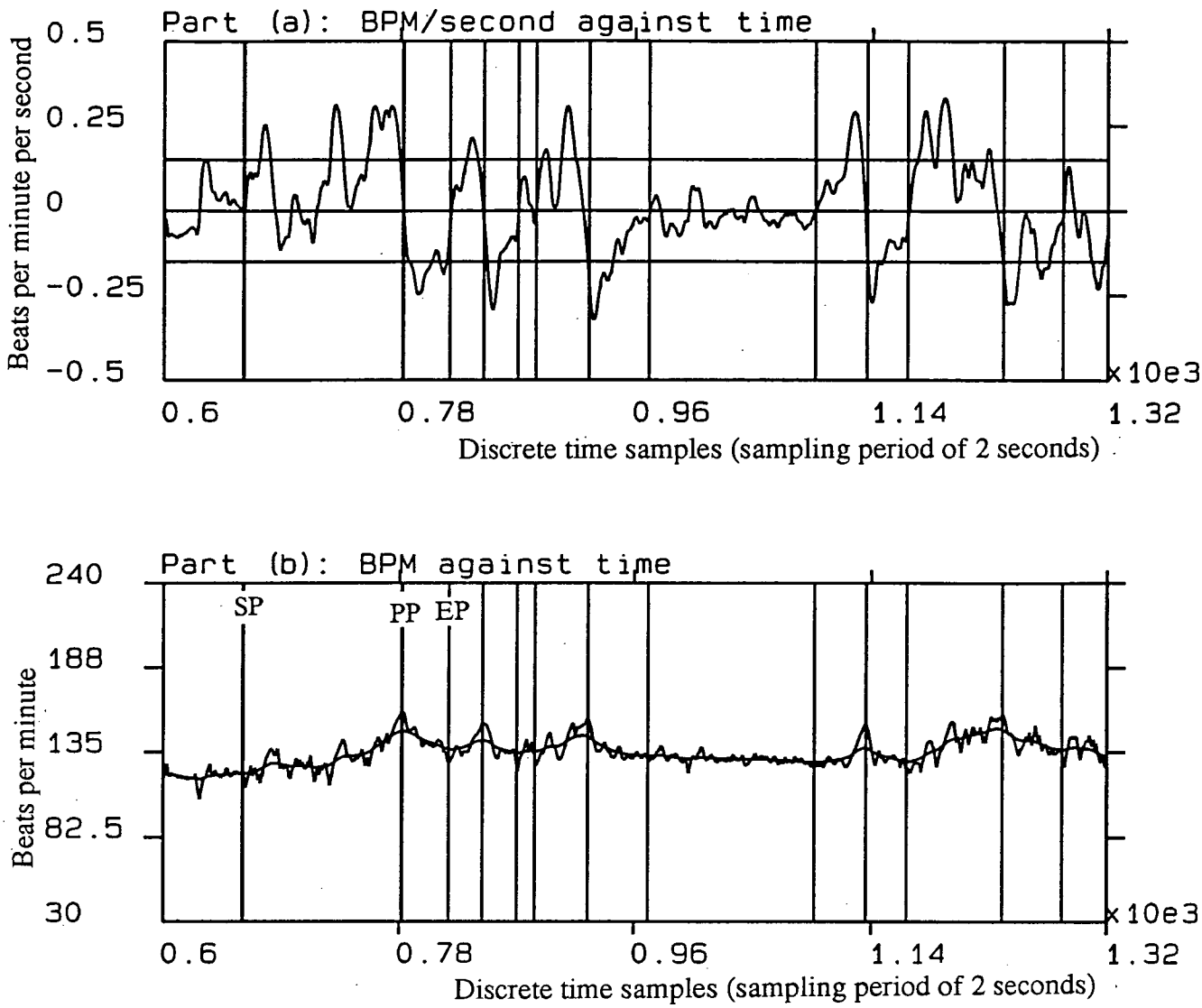


Figure 4.6: Part (a): Represents the rate of change of the baseline, shown in part (b) of the averaged FHR. The corresponding vertical lines drawn in (a) and (b) mark the points associated to starting (SP), peak (PP) and ending (EP) point of accelerations detected by the rule based routine. For example, in the case of this FHR data window shown, the routine will provide the following numerical information on each detected acceleration:

SPv: FHR value at SP; PPv: FHR value at PP; EPv: FHR value at EP; Length: duration of accelerations presented in terms of number of samples.
All the SP, PP, and EP values quoted here are presented in terms of sampling point position in the FHR data File being analyzed.

No.	SP.	SPv.(BPM)	PP.	PPv.(BPM)	EP.	EPv.(BPM)	Length.
1	661	123.226	782	159.940	818	128.333	157
2	818	128.333	844	152.428	870	127.188	52
3	884	131.164	925	156.203	971	129.645	87
4	1097	127.638	1136	151.914	1167	120.885	70
5	1167	120.885	1240	158.816	1286	134.052	119

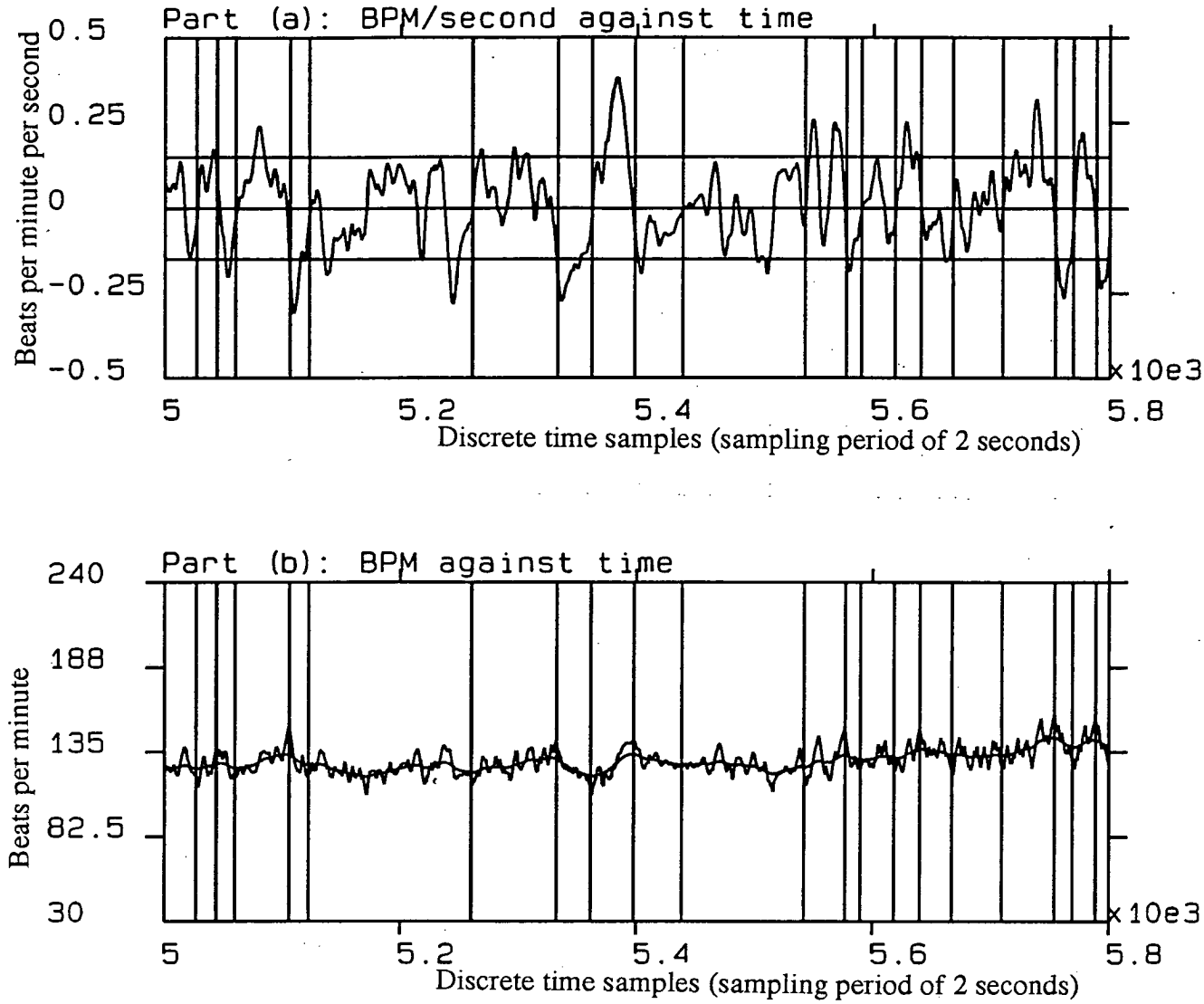


Figure 4.7: Part (a): Represents the rate of change of the baseline, shown in part (b) of the averaged FHR. The corresponding vertical lines drawn in (a) and (b) mark the points associated to starting (SP), peak (PP) and ending (EP) point of accelerations detected by the rule based routine. For example, in the case of this FHR data window shown, the routine will provide the following numerical information on each detected acceleration:

SPv: FHR value at SP; PPv: FHR value at PP; EPv: FHR value at EP; Length: duration of accelerations presented in terms of number of samples.
All the SP, PP, and EP values quoted here are presented in terms of sampling point position in the FHR data File being analyzed.

No.	SP.	SPv.(BPM)	PP.	PPv.(BPM)	EP.	EPv.(BPM)	Length.
1	5027	119.717	5044	138.225	5060	121.323	33
2	5060	121.323	5106	150.256	5122	121.075	62
3	5261	116.685	5333	142.018	5362	111.691	101
4	5362	111.691	5398	141.300	5439	124.825	77
5	5542	122.951	5577	149.614	5590	124.444	48
6	5618	126.453	5640	148.909	5667	123.793	49
7	5709	125.659	5754	157.842	5769	129.527	60

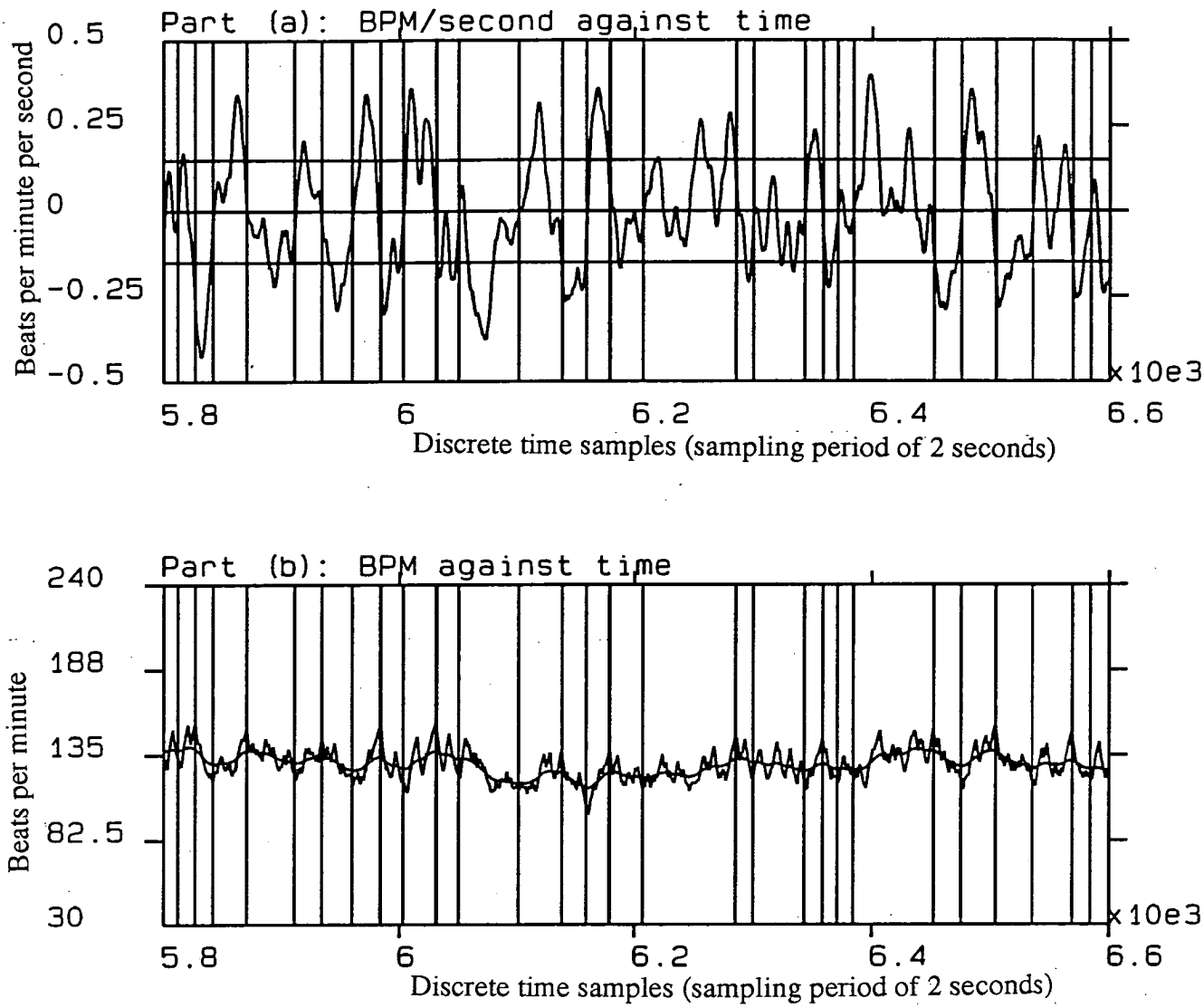


Figure 4.8: Part (a): Represents the rate of change of the baseline, shown in part (b) of the averaged FHR. The corresponding vertical lines drawn in (a) and (b) mark the points associated to starting (SP), peak (PP) and ending (EP) point of accelerations detected by the rule based routine. For example, in the case of this FHR data window shown, the routine will provide the following numerical information on each detected acceleration:

SPv: FHR value at SP; PPv: FHR value at PP; EPv: FHR value at EP; Length: duration of accelerations presented in terms of number of samples.
All the SP, PP, and EP values quoted here are presented in terms of sampling point position in the FHR data File being analyzed.

No.	SP.	SPv.(BPM)	PP.	PPv.(BPM)	EP.	EPv.(BPM)	Length.
1	5812	133.656	5827	154.918	5842	118.219	30
2	5842	118.219	5871	151.061	5911	124.065	69
3	5911	124.065	5934	143.592	5960	119.685	49
4	5960	119.685	5984	152.666	6003	118.688	43
5	6003	118.688	6031	154.547	6050	123.320	47
6	6101	113.097	6138	138.124	6159	103.626	58
7	6159	103.626	6179	136.139	6207	113.576	48
8	6207	113.576	6286	145.350	6301	120.026	94
9	6344	117.686	6359	144.883	6371	121.561	27
10	6384	122.275	6452	150.309	6475	119.400	91
11	6475	119.400	6503	153.128	6535	120.156	60
12	6535	120.156	6569	149.295	6584	121.666	49

value, peaks, returns to zero and goes negative. It then passes a heuristically found negative threshold value, peaks in the negative direction and returns back to zero again. The latter point is associated with completion of a possible FHR acceleration, where the FHR baseline (estimated trend) has returned to the steady state FHR baseline.

Finally if it also satisfies the *genuine* FHR acceleration condition mentioned earlier, then an acceleration has been found, and its numerical representation is recorded in the form shown in figures 4.3 to 4.8.

In summary The routine numerically scans forward in time through two data sets (the original FHR data, and the rate of change of the trend). It looks for a possible FHR acceleration, and then checks to see if it is a genuine FHR acceleration. If it is a genuine FHR acceleration, it proceeds by scanning back and forth in the vicinity of the detected

FHR acceleration in order to extract the relevant numerical information concerning the acceleration. The whole routine is written in the high level C-programming language and the rules used were based on the familiar IF - THEN format.

4.6 Conclusions and Discussion

In this chapter FHR variability has been analyzed by applying classical scalar statistical analysis of the detrended FHR time series. The latter was segmented into contiguous data blocks of one minute duration and their standard deviation was taken. The length of the window was short enough to have local statistical stationarity and long enough for it to be considered as a large sample, for the purpose of having unbiased standard deviation estimation.

The approach is successful in tracking the changes in FHR variability, because it bases its numerical quantification of the FHR variability on short blocks of detrended FHR data. The computed statistical values represent a time record on the true FHR fluctuations over the baseline.

The presentation of these computed standard deviation statistics, using relative frequency histogram representation allows the general status of the FHR variability in long term antepartum FHR recordings to be observed efficiently and accurately.

The information concerning the shape of these histograms can then be summarized in four moment based statistical quantities. These are, respectively, the arithmetic mean, the sample standard deviation, the coefficient of skewness and the coefficient of kurtosis of the distribution. These four parameters will be of particular value when large quantities of data are being analyzed and numerical comparative analysis between their distributions is required.

The other vital component of FHR, the FHR accelerations which are visually detected by the medics, can now be accurately detected through a ruled based structure that has been devised. The routine is consistent non-exhaustive (unlike the human expert) and can accurately scan hours of antepartum recorded FHR data for accelerations.

The detection routine not only counts the number of the FHR accelerations that have occurred but also provides detailed numerical information describing each detected FHR acceleration. This is something that the human operator can not provide through mere visual analysis of long term FHR recordings. For each detected FHR acceleration, the routine provides estimates of the starting point, the peak point, and the ending point along with their respective FHR values in BPM.

Going back to the analysis of variability, a very useful approach based on conven-

tional statistical analysis has been presented, but one thing that this type of one-dimensional statistical analysis routines lacks is the ability to actually represent the variety of random non-white variability patterns which are observed in FHR (detrended) time series. This subject will be addressed in the next chapters, where a new method of representing non-white FHR variability patterns parametrically through use of model based statistical time series analysis techniques is presented.

The need for such an analysis arises due to the incapability of the one-dimensional standard deviation statistic to discriminate between different types of non-white variability patterns. Due to the random nature (*stochastic trend*) of the data, the models that will be able to represent such patterns will be stochastic probabilistic based. In the next chapter the *time series analysis* theorem will be introduced. This is an area which specifically deals with stochastic processes and stochastic based models.

CHAPTER 5

TIME SERIES ANALYSIS OF FETAL HEART RATE VARIABILITY

5.1 Introduction

In this chapter the objective is to study the underlying stochastic (random) mechanism which gives rise to long term FHR variability, one of the important components of FHR time series. The variability with respect to the FHR baseline is not statistically white and it exhibits a variety of random (statistically dependent) patterns. As yet no successful method has been proposed to quantitatively represent these patterns. Simple one-dimensional statistical measures are not capable of sufficiently representing these patterns and are, in fact, completely inadequate.

A much more efficient statistical routine has to be applied to characterize variability patterns in FHR data. This has lead to the adoption of a *time series analysis* approach, for finding a suitable stochastic model which can represent FHR variability patterns. The assumption is that patterns observed in short, locally stationary FHR blocks can be represented numerically by the parameters of a suitable stochastic model. The main difficulty lies in finding the right (parsimonious) model.

The complete time series analysis process can be divided into three stages: (1):- *Model Identification* (or *Specification*), (2):- *Model Parametric Estimation* (or

Model Fitting), (3):- *Model Diagnostics* (or *Adequacy Checking*). The most important stage is the model *identification* or *specification* stage. This chapter has been dedicated to providing a general understanding of time series analysis and of the steps involved in identifying an adequate model for FHR time series. The aim of this modeling is to numerically represent the non-white dynamic variability patterns on a short block by block basis in terms of the parameters of the chosen model. When the medical significance of the observed variability patterns is understood, then it may be possible to associate the numerical values of the extracted parameters with particular medical conditions.

Once a tentative model has been identified, its parameters need to be estimated, and then, to check if the right model has been chosen, diagnostic checking for the adequacy of the chosen model has to be performed. Thus, Model parameter estimation (or model fitting) and model diagnostics are important subjects which will be discussed in the next chapter.

As discussed in chapter 2, FHR estimates produced by indirect (external) monitoring techniques are averaged over the duration of the observation block (2 seconds in this case) and consequently do not contain, short term beat-to-beat variability information. In this case the irregular fluctuations observed around the baseline are referred to as long term variability components of FHR. They are represented completely by the detrended FHR time series.

The fundamental concepts of time series analysis are introduced in the next section.

5.2 Time Series Analysis

A time series is an ordered sequence of observations. It is a record of the values of any fluctuating quantity measured at different points of time. Time series occur in a variety of fields, such as agriculture, weather, business and economy, engineering and medicine. In the case of this project, FHR is a discrete time series which is under analysis.

In the development of the theory of stochastic processes, an important role has been played by the study of *economic time series* [59-64]. The analysis of economic time series has been a problem of great interest to economic theorists desiring to explain the *dynamics* of economic systems and to speculators desiring to forecast prices.

There are various objectives for studying time series. These include the understanding and description of the stochastic mechanism (model) which gives rise to an observed series, the forecasting of future values, and optimal control of a system. In the case of this project, the objective falls in the first category.

The intrinsic nature of a time series is that its observations are dependent or correlated, and the order of the observations is, therefore, important. The time series $z(k)$ (where k represents a discrete time index), is said to be *deterministic* if future values of the series can be exactly described by some mathematical function. On the other hand, $z(k)$ is said to be a *statistical time series* if future values of the series can be described only in terms of a probability distribution. A statistical time series represents one particular realization of a *stochastic process*, that is, a statistical phenomenon that evolves in time according to probabilistic laws.

In order to analyze a time series $z(k)$, one must first assume a model for $z(k)$, which is completely specified except for the values of certain parameters which one proceeds

to estimate on the basis of an observed sample. To identify such a model, in time series analysis, the statistics of the time series itself must be examined, in effect allowing the data to decide the structure of the time series representation. This is the reason why the moment based statistics namely the "autocorrelation functions" and "partial autocorrelation functions", which will be discussed in the following sections are the fundamental tools in time series analysis. But first the meaning of stochastic processes and the meaning and importance of stationarity in time series analysis are discussed from a theoretical viewpoint.

5.2.1 Stochastic Processes and Stationarity

A stochastic process is a family of time indexed random variables $z(w,k)$, where w belongs to a sample space and k belongs to an index set. For a fixed k, $z(w,k)$ is a random variable. For a given w, $z(w,k)$, as a function of k, is called a sample function or realization. The population that consists of all possible realizations is called the *ensemble* in stochastic processes and time series analysis[42, 65, 66].

Thus, a time series such as an FHR time series can be looked upon as a realization or a sample function from certain stochastic process. With understanding that a stochastic process, $z(w,k)$, is a set of time indexed random variables defined on a sample space, one usually suppresses the variable w and simply writes $z(w,k)$ as $z(k)$. The process is referred to as a real valued process if it assumes only real values.

In general a real valued process is considered to be *stationary* if and only if the probability laws governing the process do not change with time; that is, the process is in statistical equilibrium. Specifically a random process $z(k)$ is said to be *strictly stationary* if the joint probability distribution of $z(k_1),z(k_2),\ldots z(k_n)$ is the same as the joint distribution of $z(k_1-\tau),z(k_2-\tau),\ldots z(k_n-\tau)$ for all choices of time points $k_1,k_2,\ldots k_n$ and

all choices of time lag τ. For such a process the mean function is defined as

$$\mu_k = E[z(k)], \tag{5.1}$$

the variance function of the process is

$$\sigma_k^2 = [E(z(k) - \mu_k)^2], \tag{5.2}$$

the covariance function between $z(k_1)$ and $z(k_2)$ is

$$\gamma(k_1,k_2) = E[(z(k_1)-\mu_{k_1})(z(k_2)-\mu_{k_2})], \tag{5.3}$$

and the correlation coefficient between $z(k_1)$ and $z(k_2)$ is

$$\rho(k_1,k_2) = \frac{\gamma(k_1,k_2)}{\sqrt{\sigma_{k_1}^2}\sqrt{\sigma_{k_2}^2}}. \tag{5.4}$$

Now for a strictly stationary process, since the distribution function is the same for all k, the mean function $\mu_k = \mu$ is a constant. Likewise $\sigma_k^2 = \sigma^2$ for all k and hence is also a constant. Moreover, since the bivariate distribution of $(z(k_1),z(k_2))$ is same as that of $(z(k_1+\tau),z(k_2+\tau))$ for any integer k_1,k_2 and τ, it will result in

$$\gamma(k_1,k_2) = \gamma(k_1+\tau,k_2+\tau), \tag{5.5}$$

and

$$\rho(k_1,k_2) = \rho(k_1+\tau,k_2+\tau). \tag{5.6}$$

Letting $k_1 = k-\tau$, and $k_2 = k$ will result in

$$\gamma(k_1,k_2) = \gamma(k-\tau,k) = \gamma(k,k+\tau) = \gamma_\tau \tag{5.7}$$

and

$$\rho(k_1,k_2) = \rho(k-\tau,k) = \rho(k,k+\tau) = \rho_\tau \tag{5.8}$$

Thus, for a strictly stationary process with the first two moment finite, the covariance and the correlation between $z(k)$ and $z(k+\tau)$ depends only on the time difference τ.

In practice it is often useful to define stationarity in a less restricted way than that described above. It is very difficult or impossible actually to verify a distribution function, particularly a joint distribution function from an observed time series. Thus, in time series analysis, the term second order stationary or weakly stationary, is used in terms of the moments of the process.

Generally as a requirement for weak stationarity, the mean is constant and its autocovariance only depends on the lag τ. No assumptions are made about higher moments than those of the second order. Note that having this, implies that the variance, as well as the mean is constant and also both of them must be finite. Finally for a *Normal* or *Gaussian* processes, second order stationarity implies strict stationarity and for processes which are very *non −Gaussian*, μ and γ_τ may not adequately describe the process.

In the case of FHR time series, there exists both non-stationarity in the first moment and the second moment. The first moment nonstationarity was removed through linear discrete filtering, but the detrended data still is nonstationary in the second moment. However, to analyze the variability pattern in the detrended data short contiguous epochs or windows of data must be considered one at a time, and under these circumstances, local weak (quasi) stationarity can be assumed[42, 60, 67-71] in the short span of the observation, for the purpose of applying the time series analysis technique.

It is clear that weak stationarity is a much weaker form of stationarity than "strictly stationary" or "stationary distribution". However, in linear time series analysis, second order stationarity is assumed to be a sufficient condition for it to be successfully

applied, and it is also an assumption which can be checked in a much more comprehensive manner. Also in this analysis, the underlying process is assumed to be Gaussian. This assumption holds totally for both strictly stationary and weakly stationary processes. Thus the type of stochastic model which has to be identified will be linear (Gaussian), and is known as autoregressive moving average (ARMA) structures. This structure will be introduced later on.

Finally, as discussed earlier, the statistics of the time series under observation dictates the form of the stochastic model (ARMA) which best describes the process. Indeed this is the essence of the time series analysis technique, which bases its identification of a tentative model on correlation information extracted from the data. In general two moment based statistics are used; autocorrelation and partial autocorrelation. These vital statistics will be thoroughly explained in the following two sections. Then the ARMA models that they are used to identify and the approach to model identification based on their sample estimates will be presented.

5.3 The Autocovariance and Autocorrelation Functions

For a stationary process $z(k)$, The covariance between $z(k)$ and $z(k+\tau)$ is defined as:

$$\gamma_\tau = Cov(z(k),z(k+\tau)) = E[(z(k)-\mu)(z(k+\tau)-\mu)], \tag{5.9}$$

and the correlation between $z(k)$ and $z(k+\tau)$ as:

$$\rho_\tau = Corr(z(k),z(k+\tau)) = \frac{Cov(z(k),z(k+\tau))}{\sqrt{Var(z(k))}\sqrt{Var(z(k+\tau))}} = \frac{\gamma_\tau}{\gamma_0}, \tag{5.10}$$

where

$$Var(z(k)) = E[(z(k)-\mu)^2] = \sigma^2, \text{ and } \mu = E[z(k)].$$

As a function of τ, γ_τ is called the autocovariance function (ACVF) and ρ_τ is called the autocorrelation function (ACF) in time series analysis since they represent the covariance and correlation between $z(k)$ and $z(k+\tau)$ from the same process, separated only by τ time lags.

Intuitively, one may interpret the ACVF(γ_τ) or ACF(ρ_τ) as a measure of the *similarity* between a realization of $z(k)$ and the same realization shifted to the left by τ units.

Assuming the process is stationary; as τ increases one would expect the correlation between $z(k)$ and $z(k+\tau)$ to decrease. If τ is large then the process will, loosely speaking, have *forgotten* at time $(k+\tau)$ the value it assumed at time (k). Consequently, one would expect both γ_τ and ρ_τ to decay to zero as $|\tau| \to \infty$. In general the rate at which ACVF or ACF decay to zero can be interpreted as a measure of the *memory* of the process.

The size of the ACVF coefficient (γ_τ) depends on the units in which $z(k)$ is measured. In time series analysis, it is more appropriate to use its standardized version; namely the ACF (ρ_τ). Generally for a stationary process the following properties stand:

1. $\gamma_0 = Var(z(k))$; $\rho_0 = 1$.

2. $|\gamma_\tau| \leq \gamma_0$; $|\rho_\tau| \leq 1$;

3. $\gamma_\tau = \gamma_{-\tau}$ and $\rho_\tau = \rho_{-\tau}$, for all τ, i.e., γ_τ and ρ_τ are even functions and hence symmetric about the time origin, $\tau = 0$. This follows from the fact that the time difference between $z(k)$ and $z(k+\tau)$ and $z(k)$ and $z(k-\tau)$ are the same. Therefore, the ACF (or ACVF) is usually plotted only for the nonnegative lags such as the ones shown in part (b) of figures 5.1 to 5.15. Usually such plots are known as

correlograms.

Finally, in order to calculate exactly the ACF or ACVF values of a process, the ensemble of all possible realizations must be known. Alternatively, they can be estimated if multiple independent realizations are available. But, in most situations, such as with FHR time series, it is virtually impossible to obtain multiple realizations. In most cases time series (like FHR) constitute only a single realization. This makes it impossible to calculate any ensemble statistical average values. However, since stationarity in the process is a pre-condition for such types of correlation analysis, there is a natural alternative of replacing ensemble averages by time averages. In the following sections the sample estimation technique for ACVF and ACF will be presented

5.3.1 Sample Estimation of Correlation Statistics

In general the ACF coefficients are computed by first calculating the series of ACVF coefficients (γ_τ) first. The sample autocovariance formula is defined as:

$$\hat{\gamma}_\tau = \frac{1}{N}\sum_{k=1}^{N-\tau}(z(k)-\bar{z})(z(k+\tau)-\bar{z}), \text{ where } \bar{z}=\sum_{k=1}^{N}\frac{z(k)}{N}. \tag{5.11}$$

There is also a slightly modified version of above equation, which results in slightly less biased estimation; this being:

$$\hat{\gamma}_\tau = \frac{1}{N-\tau}\sum_{k=1}^{N-\tau}(z(k)-\bar{z})(z(k+\tau)-\bar{z}), \tag{5.12}$$

For the purpose of computing the standardized sample autocorrelation function (SAC), formula (5.11) is generally used, and since $\hat{\rho}_\tau = \frac{\hat{\gamma}_\tau}{\hat{\gamma}_0}$, it results in:

$$\hat{\rho}_\tau = \frac{\sum_{k=1}^{N-\tau}(z(k)-\bar{z})(z(k+\tau)-\bar{z})}{\sum_{k=1}^{N}(z(k)-\bar{z})^2}, \tau = 0,1,2,3,.. \tag{5.13}$$

Now for a stationary Gaussian process, Bartlett[42, 62, 72-74] has shown that for $\tau > 0$, and $\tau+\zeta > 0$,

$$Cov(\hat{\rho}_\tau,\hat{\rho}_{\tau+\zeta}) \approx \frac{1}{N}\sum_{i=-\infty}^{\infty}(\rho_i\rho_{i+\zeta}+\rho_{i+\tau+\zeta}\rho_{i-\tau}-2\rho_\tau\rho_i\rho_{i-\tau-\zeta}-2\rho_{\tau+\zeta}\rho_i\rho_{i-\tau}+2\rho_\tau\rho_{\tau+\zeta}\rho_i^2) \quad (5.14)$$

For large N, $\hat{\rho}_\tau$ is approximately normally distributed with mean ρ_τ and variance:

$$Var(\hat{\rho}_\tau) \approx \frac{1}{N}\sum_{i=-\infty}^{\infty}(\rho_i^2+\rho_{i+\tau}\rho_{i-\tau}-4\rho_\tau\rho_i\rho_{i-\tau}+2\rho_\tau^2\rho_i^2);\ m = \tau-1. \quad (5.15)$$

For processes in which $\rho_\tau = 0$ for $\tau > m$, Bartlett's approximation shown in the above equation becomes:

$$Var(\hat{\rho}_\tau) \approx \frac{1}{N}(1+2\rho_1^2+2\rho_2^2+\ldots+2\rho_m^2);\ m = \tau-1. \quad (5.16)$$

In a practical situation, $\rho_i (i = 1,2,\ldots,m)$ are unknown and are replaced by their sample estimates $\hat{\rho}_i$; which results in the following large-lag standard error for $\hat{\rho}_\tau$;

$$S_{\hat{\rho}_\tau} = \sqrt{\frac{1}{N}(1+2\hat{\rho}_1^2+\ldots+2\hat{\rho}_m^2)}. \quad (5.17)$$

Now if the process $z(k)$ represents an independent and identically distributed random variables with a constant mean (white noise process), then the above equation will be further simplified and the standard error of this process becomes simply:

$$S_{\hat{\rho}_\tau} = \sqrt{\frac{1}{N}}. \quad (5.18)$$

Finding the standard error (or standard deviation) of the SAC statistics will allow us to check if statistically the lag of interest is zero or not, and in general equations (5.17) and (5.18) are used regularly for this purpose.

5.4 The Partial Autocorrelation Function

Another important characteristic of interest which is used extensively in time series analysis for model identification, is the correlation between $z(k)$ and $z(k-\tau)$ after their mutual linear dependency on the intervening random variables $z(k-1), z(k-2), \ldots,$ and $z(k-\tau+1)$ has been removed[60, 62, 63, 73-75]. For stationary time series, this correlation coefficient is called "partial autocorrelation" at lag τ and here will be denoted by $\phi_{\tau\tau}$. As a function of τ, $\phi_{\tau\tau}$ is usually known as the partial autocorrelation function(PACF).

In general the PACF lags($\phi_{\tau\tau}$) can be defined in a closed form solution in terms of the ACF lags(ρ_τ)[60, 62, 74-76]. The clearest way to explain this relation is through use of the following regression model:

$$z(k+\tau) = \phi_{\tau 1} z(k+\tau-1) + \phi_{\tau 2} z(k+\tau-2) + \ldots + \phi_{\tau\tau} z(k) + e(k+\tau), \tag{5.19}$$

where $\phi_{\tau i}$ denotes the i-th regression parameter and $e(k+\tau)$ is a normal error term uncorrelated with $z(k+\tau-j)$ for $j \geq 1$. Multiplying $z(k+\tau-j)$ on both sides of equation (5.19) and taking the expectation and dividing by γ_0 results in the following expression:

$$\rho_j = \phi_{\tau 1}\rho_{j-1} + \phi_{\tau 2}\rho_{j-2} + \ldots + \phi_{\tau\tau}\rho_{j-\tau}. \quad j = 1,2,\ldots,\tau \tag{5.20}$$

This is the well known structure of Yule-Walker equations. More explicitly for $j = 1,2,\ldots\tau$ following system of equations are derived:

$$\begin{aligned} \rho_1 &= \phi_{\tau 1}\rho_0 + \phi_{\tau 2}\rho_1 + \ldots + \phi_{\tau\tau}\rho_{\tau-1} \\ \rho_2 &= \phi_{\tau 1}\rho_1 + \phi_{\tau 2}\rho_0 + \ldots + \phi_{\tau\tau}\rho_{\tau-2} \\ &\ldots \\ &\ldots \end{aligned} \tag{5.21}$$

$$\rho_\tau = \phi_{\tau 1}\rho_{\tau-1} + \phi_{\tau 2}\rho_{\tau-2} + \ldots + \phi_{\tau\tau}\rho_0$$

Assuming the $\rho_1, \rho_2, \ldots, \rho_\tau$ values are already known, the above system of equations can then be solved for the unknowns $\phi_{\tau 1}, \phi_{\tau 2}, \ldots, \phi_{\tau\tau}$ and afterwards all but $\phi_{\tau\tau}$ coefficients may be discarded.

To solve the equation (5.20) involves extensive matrix determinant analysis. It can be shown through the use of *Cramer's rule* [77], that $\phi_{\tau\tau}$ can be represented successively for increasing τ in terms of following determinant based structures:

$$\phi_{11} = \rho_1 \tag{5.22}$$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}$$

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}$$

$$\phi_{\tau\tau} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{\tau-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{\tau-3} & \rho_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \rho_{\tau-1} & \rho_{\tau-2} & \rho_{\tau-3} & \cdots & \rho_1 & \rho_\tau \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{\tau-2} & \rho_{\tau-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{\tau-3} & \rho_{\tau-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \rho_{\tau-1} & \rho_{\tau-2} & \rho_{\tau-3} & \cdots & \rho_1 & 1 \end{vmatrix}}$$

Thus the partial autocorrelation between $z(k)$ and $z(k+\tau)$ can also be obtained as the regression coefficient with $z(k)$ when regressing $z(k+\tau)$ on its τ lagged variables $z(k+\tau-1), z(k+\tau-2), \ldots$, and $z(k)$ as in equation (5.19).

5.4.1 The Sample Partial Autocorrelation Function

For an observed time series, one needs to be able to estimate the PACF. Given the system of equations in (5.21), an obvious method is to estimate the ρ_τ's by $\hat{\rho}_\tau$'s and then solve the resulting linear set of equations for $\tau = 1,2,\ldots$ to get sample estimates of $\phi_{\tau\tau}$. This estimated function can then be called the sample PACF or SPAC for short and be denoted by $\hat{\phi}_{\tau\tau}$.

The main problem encountered in solving equation (5.21) is in calculating complicated matrix determinants for large τ, as was shown in equations (5.22). However, this problem is remedied by an efficient recursive solution given by Levinson in (1947) and later on by Durbin in (1960)[78, 79]. Since then this recursive technique have come to be known as the Levinson-Durbin algorithm, and is used for solution of the Yule-Walker equations. They showed independently; starting with $\phi_{11} = \rho_1$ the equations (5.22) can be solved by the recursion:

$$\phi_{\tau\tau} = \frac{\rho_\tau - \sum_{j=1}^{\tau-1} \phi_{\tau-1,j}\,\rho_{\tau-j}}{1 - \sum_{j=1}^{\tau-1} \phi_{\tau-1,j}\,\rho_j}, \tag{5.23}$$

where

$$\phi_{\tau j} = \phi_{\tau-1,j} - \phi_{\tau\tau}\phi_{\tau-1,\tau-j}, \quad \text{for } j = 1,2,\ldots,\tau-1.$$

Thus using equation (5.23), it is possible to numerically calculate as many values for $\phi_{\tau\tau}$ as desired. As stated these equations give only the theoretical PACF, but by

replacing the ρ_τ by $\hat{\rho}_\tau$, in equation (5.23), the SPAC $\hat{\phi}_{\tau\tau}$ can be obtained.

Finally, in the case of white noise process, the theoretical PACF lag values will all be zero, but this is not so in the case of SPAC lags, since they are sample estimates. The SPAC lags for a Gaussian white noise process will be quantities very close to zero. It will be of much value to know the standard error or standard deviation of $\hat{\phi}_{\tau\tau}$ statistics. In general, under the hypothesis that the underlying process is a white noise sequence, the standard error of $\hat{\phi}_{\tau\tau}$ can be approximated[42, 59, 62, 63, 73, 75, 80-82] by

$$S_{\hat{\phi}_{\tau\tau}} = \sqrt{\frac{1}{N}}. \tag{5.24}$$

Hence, if the underlying process is truly white Gaussian noise, the computed SPAC lags should be distributed approximately normally about zero, and they should lie within plus or minus two standard errors of $2/\sqrt{N}$, N being the number of samples in the observation. In general the standard error is computed using this equation and can be used to test the statistical significance of estimated partial autocorrelation coefficients.

Finally, the standard error introduced earlier for SAC ($S_{\hat{\rho}_\tau}$) is the same as the one for SPAC when the respective spikes are suppose to be insignificant and showing lack of correlation. In the next section the general type of stochastic models used in time series analysis will be introduced and their relation with ACF and PACF will be clarified as a prelude to the specification or identification routine which makes use of SAC and SPAC.

5.5 Stochastic Time Series Models

Given that we are always limited by a finite number of available observations in time series analysis problems, it is intuitively logical also to have a finite order parametric model to describe a time series process. In this section moving average (MA), autoregressive (AR) and autoregressive moving average (ARMA) time series models will be introduced. After describing the characteristics of each process in terms of ACF and PACF, the model identification strategy which is based on SAC and SPAC will be introduced and afterwards it will be applied to real FHR data.

5.5.1 Moving Average Models

The moving average structure coincides exactly with the general infinite order linear stochastic process[62], represented as a weighted linear combination of the present and past terms of a white noise process $e(k)$:

$$z(k) = e(k)+\psi_1 e(k-1)+\psi_2 e(k-2)+\psi_3 e(k-3)+\ldots \tag{5.25}$$

However, in terms of moving average, if only a finite number of ψ terms are nonzero, i.e., $\psi_1 = -\theta_1$, $\psi_2 = -\theta_2,\ldots,$ $\psi_q = -\theta_q$ and $\psi_k = 0$ for $k>q$, then the resulting structure is called a moving average process of order q, usually represented as MA(q) and given by

$$z(k) = e(k)+\theta_1 e(k-1)+\theta_2 e(k-2)+\ldots+\theta_q e(k-q) \tag{5.26}$$

It can be proven through mathematical analysis of $Corr(z(k),z(k-q))$ for a MA(q) process[60, 62, 73-75, 81, 82]; that the theoretical ACF lag will have significant spike values up to the q-th lag and abruptly fall exactly to zero for lags beyond q. Simultaneously its corresponding theoretical PACF will tail off as a mixture of exponential decays and/or damped sine waves depending on the nature of the roots of the MA(q)

characteristic equation.

This phenomenon can be manipulated to identify MA(q) type structures. Given that there are enough samples available from an observation (window), if the process is genuinely an MA(q); then the SAC and SPAC will behave in a similar manner to the theoretical ACF and PACF. Of course, since the SAC and SPAC are sample estimates; tests of significance of the computed correlation spikes are necessary. These will be introduced in the next section. Finally, for general interest, Slutzky in 1927 and later on Wold in 1938 [62, 74] were the first people to consider moving average models.

5.5.2 Autoregressive Models

Autoregressive processes are, as their name implies, regressions on themselves. Specifically a p-th order *autoregressive process* $z(k)$ satisfies the following equation:

$$z(k) = \phi_1 z(k-1) + \phi_2 z(k-2) + \ldots + \phi_p z(k-p) + e(k) \tag{5.27}$$

The current value of such a series is a linear combination of the p most recent past values of itself plus an *innovation* term $e(k)$, which incorporates everything new in the series at time k that is not explained by the past values. Thus $e(k)$ is independent of the $z(k-1), z(k-2), \ldots$ past values. Essentially, $e(k)$ is a zero mean white noise process.

AR processes are useful in describing situations in which the present value of a time series depends on its preceding values plus a random shock. From a historical point of view, it would be of value to mention that *Yule* (1927)[42, 62] carried out the original work on AR processes. He used an AR process to describe the phenomena of sunspot numbers and the behaviour of a simple pendulum.

Similarly as for the MA(q) structure, there is a similar pattern of behaviour in regards to the theoretical ACF and PACF for AR(p), but in an opposite manner. It can be proven mathematically that the ACF for an AR(p) structure behaves in the same manner as a PACF for an MA(q) structure in that it tails off to zero as a mixture of exponential and/or damped sine waves. Similarly, the PACF for an AR(p) process behaves in the same manner as the ACF for an MA(q) process in that, PACF for an AR(p) process will cut off to absolute zero at lags beyond p and not including p itself.

The theoretical analysis could begin by multiplying the equation for an AR(p) structure by $z(k-\tau)$ and taking its expectation. This will result in a recursive solution for the theoretical ACF of an AR(p) structure, which turns out to be the Yule-Walker equations introduced earlier (equation (5.20)). Notice that the p-th coefficient of an AR(p) model, ϕ_p, is equivalent to the pth-lag PACF value $\phi_{\tau\tau}$.

Now, by using the fact that for an AR(p) process the ACF $\rho_\tau = \phi_1\rho_{\tau-1} + \phi_2\rho_{\tau-2} \ldots + \phi_p\rho_{\tau-p}$ for $\tau>0$, one can easily see that when $\tau>p$, the last column of the determinant in the numerator of $\phi_{\tau\tau}$ in equation (5.22) can be written as a linear combination of previous columns of the same matrix. Hence, the PACF $\phi_{\tau\tau}$ will vanish after lag p. This is a very useful property in identification of AR(p) based processes.

Finally, in practical situations the analysis will be based on sample estimates and the SAC and SPAC will behave in a similar manner as ACF and PACF for an AR(p) process, given there is enough samples. Again to pick out the significant lag estimates of SAC and SPAC, a statistical test of significance is vitally important. This will be discussed in detail later on.

5.5.3 Autoregressive Moving Average Processes

In some time series, the process may have both AR and MA characteristics. This requires an *autoregressive moving average* model of order p and q (ARMA(p,q)) to be used in order to be able to adhere with the principle of *parsimony* [60, 62, 63]. In general an ARMA(p,q) process is represented by the following equation

$$z(k) = \phi_1 z(k-1) + \phi_2 z(k-2) + \ldots + \phi_p z(k-p) + \ldots + e(k) \\ -\theta_1 e(k-1) - \theta_2 e(k-2) - \ldots - \theta_q e(k-q). \tag{5.28}$$

Since both AR and MA characteristics are present in a process, neither its ACF nor its PACF will show any sharp cutoff to zero after q or p lag respectively. In fact, in ARMA type processes, these two statistics will both gradually decay to zero as a mixture of exponential or damped sine waves. In general the ACF tails off after lag (q-p) and the PACF tails off after lag (p-q). Again, given enough sample points from such a process, the SAC and SPAC will behave in a similar manner.

5.6 Model Identification or Specification Criteria

In the previous section the relationships between the ACF and the PACF with AR,MA and ARMA structures were introduced. In this section these behaviours will be used to produce a set of rules for identification or specification of an adequate model for the underlying process which governs the behaviour of a time series.

In practice, the theoretical ACF and PACF are unknown and, for a given observed time series of N samples, they must be estimated by the SAC ($\hat{\rho}_\tau$), and SPAC ($\hat{\phi}_{\tau\tau}$) discussed earlier. Thus the goal is to match patterns in the SAC and SPAC with known patterns of ACF and PACF for the ARMA models. In order to be able to achieve this objective more effectively, statistical tests of significance of the computed

correlation spikes are also necessary, in order to discriminate between significant and insignificant (theoretically zero) spikes.

Ideally a minimum of $N = 50$ observations is required, to identify and build a reasonable MA, AR or ARMA model. The number of SAC and SPAC lags to be calculated should be about $N/4$[60, 62, 74, 81-83], although occasionally for data of good quality one may be able to identify an adequate model with a smaller sample size, N.

In general, given there are enough samples available, the set of rules for tentatively identifying the correct stochastic models are:

1. For an MA(q) structure, the SAC will cut-off after lag q and the SPAC will die down in a damped exponential fashion with or without oscillation.

2. For an AR(p) model, the SAC dies down in the same way as the SPAC for MA(q). The SPAC will have significant spikes at lags 1,2,...,p, and will cut off after lag p to insignificant values (theoretically zero).

3. For a mixed ARMA(p,q) model, both the SAC and the SPAC will die down in a damped exponential fashion with or without oscillation. The SAC will tail off after lag (q-p) and the SPAC will after lag (p-q).

4. When the SAC decays very slowly (and usually linearly), is an indication of nonstationarity in the series (mainly due to trend). The nonstationary trend needs to be removed before any successful time series analysis can be applied.

The rate of exponential (sinusoidal) decay towards zero for the SAC and the SPAC should be rapid. In practice they will be considered to have decayed rapidly if and only if their absolute t-values (explained earlier) fall below roughly 1.6 by about lag 5 or 6[83]. This is especially important in the case of the SAC, in order to detect the

possibility of mean nonstationarity in the data.

It is useful and interesting to note that a strong duality exists between the AR and MA models in terms of their ACFs and PACFs. This identification strategy of the ARMA models is also known as *Univariate Box–Jenkins* model identification method, since the first comprehensive book that was ever written on this topic was by these authors [60].

In cases where the time series is noisy, i.e. the original process has been contaminated by external noise, if the underlying process were pure AR, the SAC and the SPAC will behave as if they were looking at an ARMA process. This is due to the presence of noise in the series. Therefore it is imperative for first time identification of a time series, to collect very good data, which is not unduly corrupted by noise.

Finally, It would be of value to introduce in this section before going for model specification for long term FHR variability, the concepts of *parsimony* and *identifiability*. These are important terms which are used to answer the main question of what constitutes a *good* model in statistical time series analysis.

5.6.1 Parsimony and Identifiability

The mechanisms by which the FHR time series is generated are complex. When a model is constructed, the objective is not to produce an accurate description of these real world mechanisms. On the contrary, the aim is to simplify the underlying process in such a way that only the essential features are retained.

Because the process is unknown, one would never know if one has selected a model that is essentially the same as the true generating process. All one can do is to select a model that seems adequate in light of the available data.

Essentially the model used should require the smallest possible number of parameters that will adequately represent the data. This approach, advocating simplicity as far as possible has come to be known as the *principle of parsimony* in time series analysis[60, 62, 63, 71, 81-83]. Albert Einstein was quoted in Cryer[62] as remarking that *everything should be made as simple as possible but not simpler*.

A parsimonious model fits the available data adequately without using any unnecessary coefficients. One is happy to find a model which can approximate the true process, as long as the model explains the behaviour of the available realization in a parsimonious and statistically adequate manner. The importance of the principle of parsimony cannot be overemphasized.

The other important factor is the principle of *identifiability*. This concept is fundamental to any statistical model. The main point to note is that if a model is not identifiable, the estimates, cannot, in general, be interpreted in any meaningful way. In such a situation more than one set of parameter values is consistent with the data. Although identifiability is a precisely defined statistical concept, it is related to parsimony. As a general rule, the more parsimonious a model, the less likely it is to suffer from problems of identifiability.

5.6.2 Statistical Tests of Significance

In classical scalar statistics one of the main aims is to know as much as possible about the population characteristics of a process. However, obtaining all relevant information is usually impossible or too costly. In practice, the information concerning the population is derived using a sample (realization), together with some probability concepts, formulas and theories from mathematical statistics. The sample is used to *estimate* a characteristic of the population or to *test a hypothesis* about the population.

This kind of inductive reasoning is known as *statistical inference*.

In time series analysis statistical inference is employed at all of its three stages. At the specification stage the estimated SAC and SPAC from a realization are used to help tentatively select a parsimonious model to represent the unknown process that generated the realization.

The equations (5.17) and (5.25) introduced earlier, for standard error of the SAC ($S_{\hat{\rho}_\tau}$) and SPAC ($S_{\hat{\phi}_{\tau\tau}}$) respectively are the main equations used by the time series analyst to test the estimated correlation spikes for significance. In general, SAC lag values will be considered as insignificant if and only if they lie within the $\pm 2S_{\hat{\rho}_\tau}$ interval. Similarly for SPAC lag values they will be considered as being statistically insignificant if they lie within the $\pm 2S_{\hat{\phi}_{\tau\tau}}$ interval.

In order to make the detection more comprehensive, the *t −statistics* [62, 63, 83] of the estimated correlation spikes are used to test the *null hypothesis* [49], $H_0 : \rho_\tau = 0$ for $\tau = 1,2,3,....$ This null hypothesis is tested by finding out how far away the sample statistics $\hat{\rho}_\tau$ and $\hat{\phi}_{\tau\tau}$ are from the hypothesized value $\rho_\tau = 0$ and $\phi_{\tau\tau} = 0$ respectively.

Now "how far" in this case is a t-statistic equal to a certain number of estimated standard errors. Therefore for $\hat{\rho}_\tau$, an approximate t-statistic is defined in this way:

$$t_{\hat{\rho}_\tau} = \frac{\hat{\rho}_\tau - \rho_\tau}{S_{\hat{\rho}_\tau}} = \frac{\hat{\rho}_\tau}{S_{\hat{\rho}_\tau}}; \; since \; H_0 : \rho_\tau = 0 \tag{5.29}$$

Equivalently for $\hat{\phi}_{\tau\tau}$, there is the following t-statistic approximation:

$$t_{\hat{\phi}_{\tau\tau}} = \frac{\hat{\phi}_{\tau\tau} - \phi_{\tau\tau}}{S_{\hat{\phi}_{\tau\tau}}} = \frac{\hat{\phi}_{\tau\tau}}{S_{\hat{\phi}_{\tau\tau}}}; \; since \; H_0 : \phi_{\tau\tau} = 0. \tag{5.30}$$

In general for a sample correlation spike ($\hat{\rho}_\tau$ or $\hat{\phi}_{\tau\tau}$) to be considered as statistically significant, its corresponding t-statistic must have an absolute value greater than 2,

otherwise its significance can be rejected[63, 83].

5.6.3 Other Identification Approaches

There are a number of other identification methods. One of the most studied is Akaike's (1973) information criteria (AIC)[84]. Using this approach involves selecting a few probable model structures, fitting them to the data available and seeing which one of them minimizes:

$$AIC(k) = -2 \ln (\textit{maximum likelihood}) + 2k. \tag{5.31}$$

where k is the total number of AR and MA parameters in the model. Maximum likelihood estimation will be discussed in chapter 6. The addition of the term $2k$ serves as a "penalty function" to avoid consideration of models with too many parameters and to help ensure the selection of parsimonious models.

The AIC technique may be of value when the underlying mechanism is found to have correlation characteristics equivalent to mixed ARMA type structures, or in situations where there may be several adequate models that can be used to represent a given data set, and there might be the need to know which is the best one.

However, Shibata (1976) has shown that the AIC criterion tends to overestimate the order of models which are exclusively autoregressive[85]. A similar observation was also made by Sneek (1984), where he suggested that there is a tendency for the AIC to choose a model which is overparameterized[86].

Parzen (1974) introduced his "criterion autoregressive transfer function" (CAT), to be optimized[87]. Cleveland (1972) defined the concept of *inverse autocorrelations* [88]. These are just a few of the many approaches for identification.

Generally speaking, model identification is always tentative. All these identification routines should be looked upon only as guides. Having chosen a model, then the parameters of the model must be estimated as efficiently as possible and then the implications of that model need to be considered. If the model is inappropriate for some reason, this must be detected at this parametric estimation stage, allowing a new, better model to be specified.

5.7 Model Identification of Actual Time Series

In this section the theory will be put into practice and the procedure that was explained in detail for model identification will be applied to actual short locally stationary blocks of detrended FHR data. The blocks were each of two minutes duration, corresponding to 60 data samples and the number of lags computed was 15, equivalent to a quarter of the number of samples being observed.

The type of model that could have arisen, might have been pure AR, pure MA, or a mixture of the two. All of these possible eventualities were taken into consideration.

However, when the identification procedure was applied to short (locally stationary) data windows, it was clearly observed that the underlying generating mechanism for long term FHR variability exhibited exclusively pure AR type statistical correlation characteristics, as will be shown in the following identification examples.

Self explanatory figures 5.1 to 5.15 show the representative SAC and SPAC type of patterns two minute observation windows obtained from a detrended long term phon-cardiographic FHR recording. In these graphical examples the two minute windows of the data under observation are boxed by a vertical line (0 to 60 samples). Each two minute realization is shown together with the corresponding SAC and SPAC functions. The plus and minus 2 standard error limits are also shown - dot/dashed lines -

and - as are the computed tSAC and tSPAC values.

Inspection of figures 5.1 to 5.15 will reveal that, the computed SAC values, for all the data windows shown, either decayed exponentially towards zero or showed damped sine waves pattern, and the computed SPAC values showed significant correlation spikes either at lag one on its own, or at lags one and two. The values for the SAC and SPAC lags at higher lag values were all insignificant, since they all fell within their ± 2 standard deviation 95% confidence limits set for zero correlation value (insignificant correlation spikes).

Such correlation signature patterns undoubtedly confirms the pure autoregressive characteristics of the underlying mechanism responsible for the long term variability patterns. They have also tentatively identified both AR(1) and AR(2) as suitable modeling candidates to represent short locally stationary windows of detrended data.

However, in this study the primary objective is to find and make use of a single model, which can statistically adequately fit to all the short locally stationary segments. On an intuitive basis, one can see that the second order autoregressive structure is the one which can satisfy this condition. In summary, the AR(2) model is the tentative identified structure for parsimonious parametric representation of long term FHR variability patterns. Its statistical adequacy, in terms of being the right parsimonious model, will be investigated further in the next chapter.

5.8 Summary and Discussion

In time series analysis, the most crucial steps are to identify and build a model based on the available data. This requires a good understanding of the processes discussed in

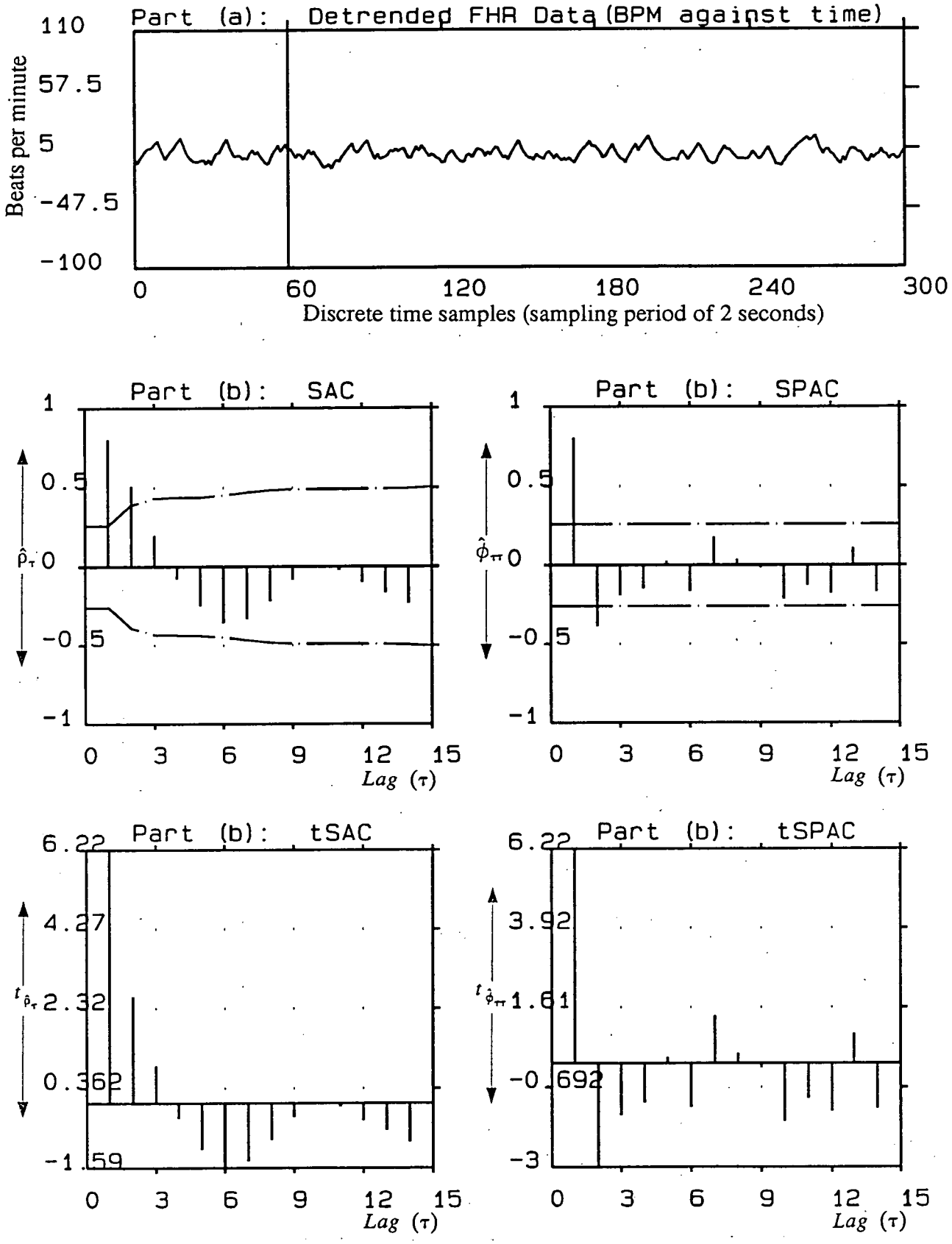


Figure 5.1: Part (a): Detrended averaged FHR (indirect phonocardiographic).
Part (b): SAC, SPAC, and t-statistics of SAC and SPAC of boxed segment shown in part (a); indicating an AR(2) process. The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.

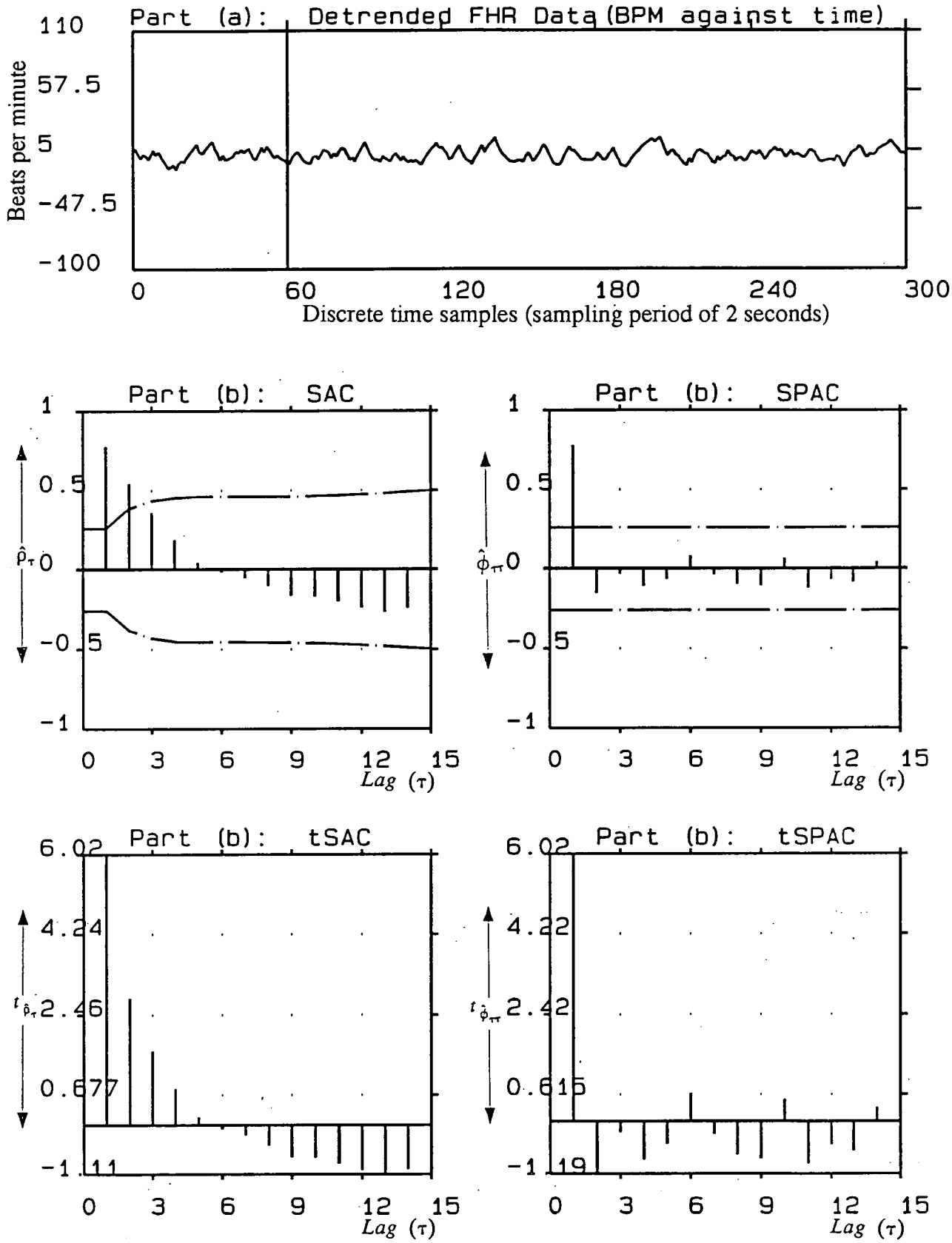


Figure 5.2: Part (a): Detrended averaged FHR (indirect phonocardiographic).
Part (b): SAC, SPAC, and t-statistics of SAC and SPAC of boxed segment shown in part (a); indicating an AR(1) process. The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.

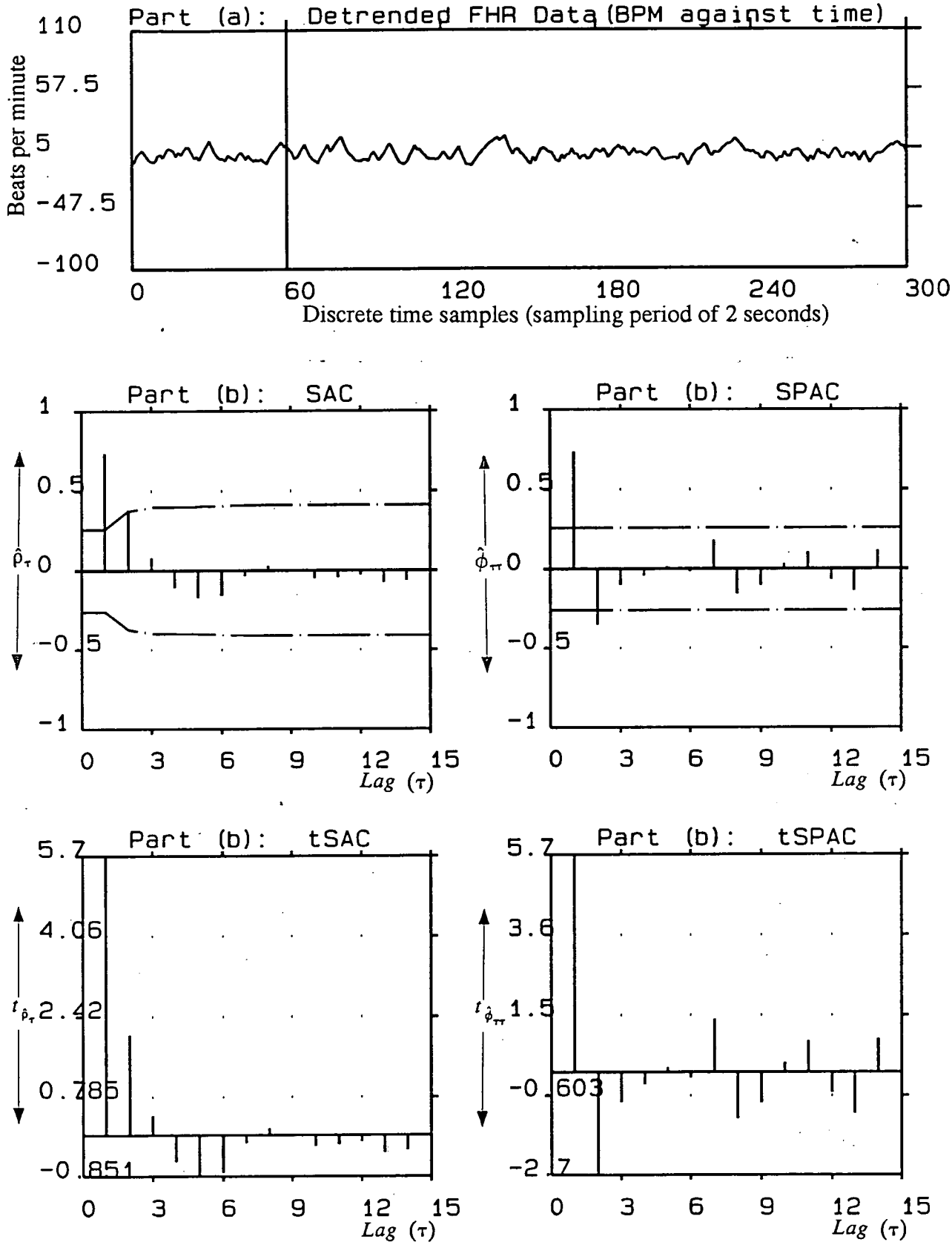


Figure 5.3: Part (a): Detrended averaged FHR (indirect phonocardiographic).
Part (b): SAC, SPAC, and t-statistics of SAC and SPAC of boxed segment shown in part (a); indicating an AR(2) process. The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.

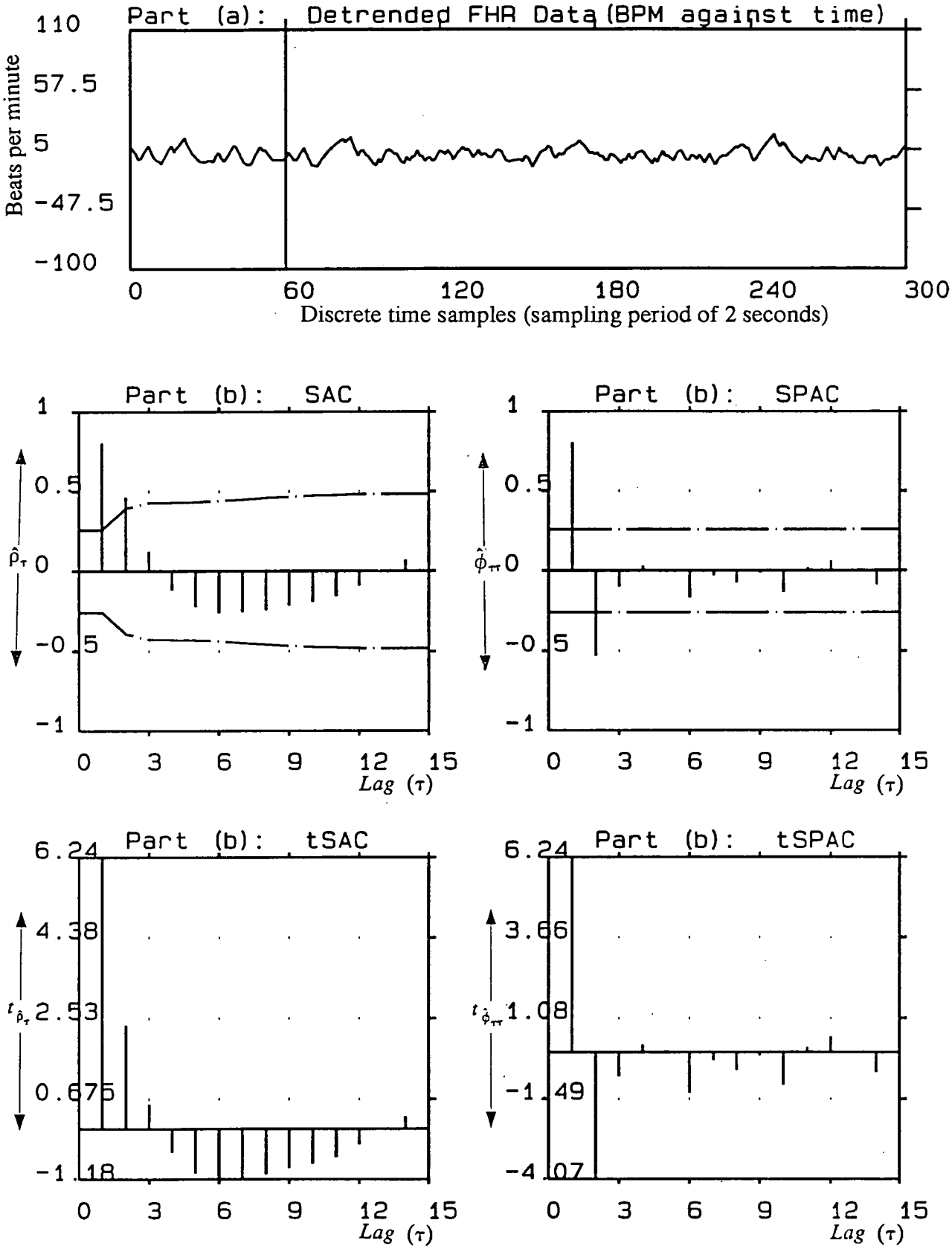


Figure 5.4: Part (a): Detrended averaged FHR (indirect phonocardiographic).
Part (b): SAC, SPAC, and t-statistics of SAC and SPAC of boxed segment shown in part (a); indicating an AR(2) process. The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.

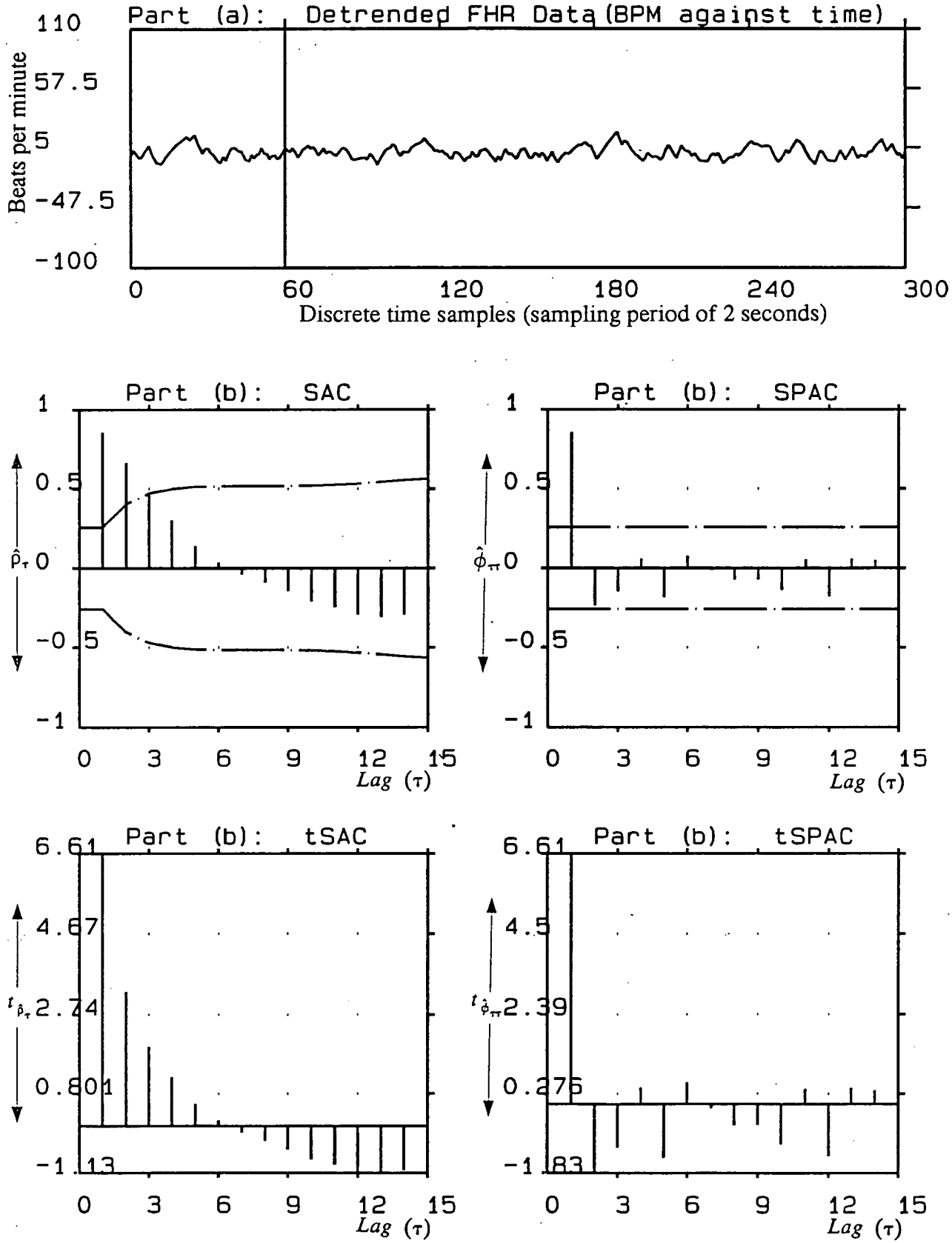


Figure 5.5: Part (a): Detrended averaged FHR (indirect phonocardiographic).
Part (b): SAC, SPAC, and t-statistics of SAC and SPAC of boxed segment shown in part (a); indicating an AR(1) process. The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.

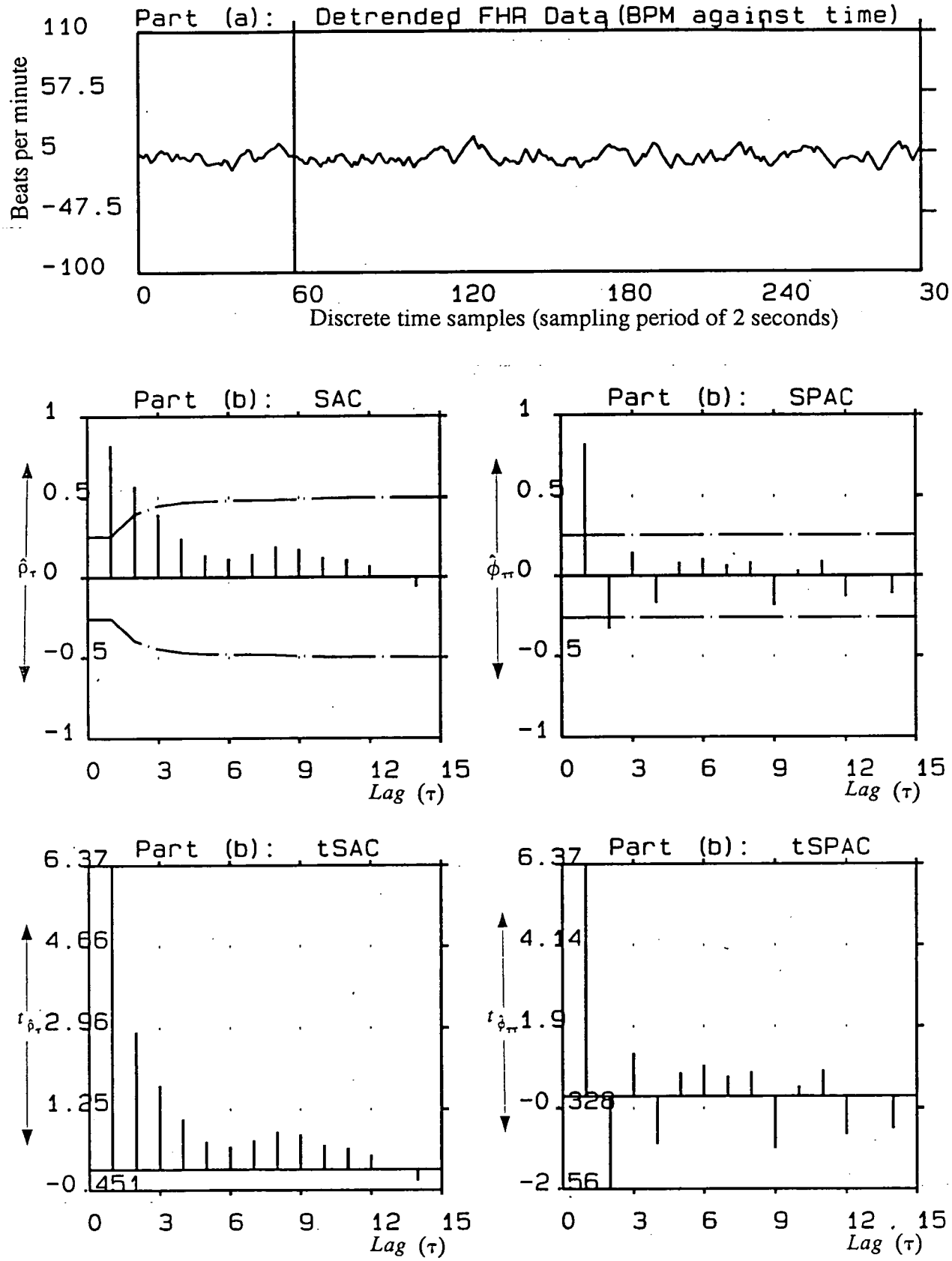


Figure 5.6: Part (a): Detrended averaged FHR (indirect phonocardiographic).
Part (b): SAC, SPAC, and t-statistics of SAC and SPAC of boxed segment shown in part (a); indicating an AR(2) process. The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.

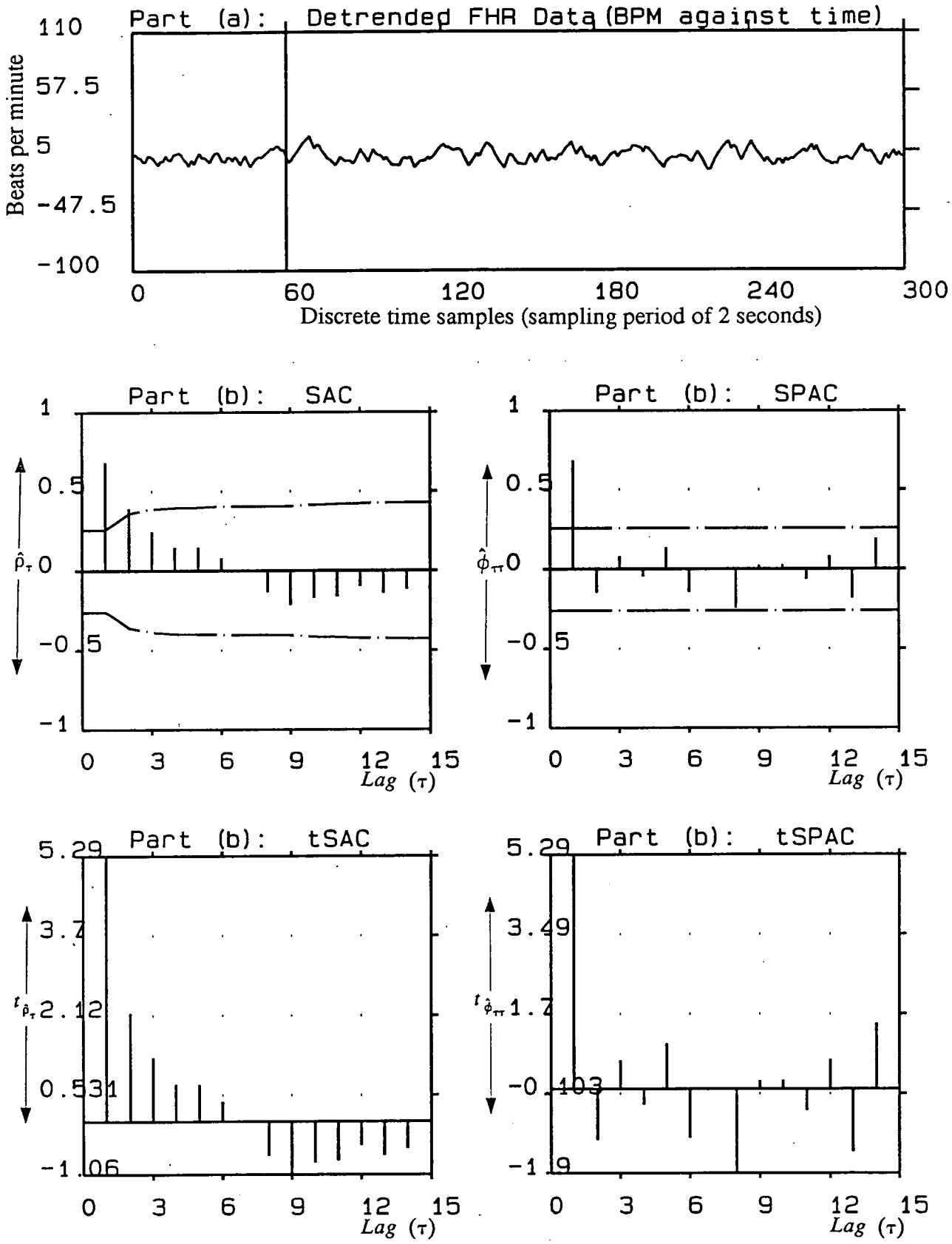


Figure 5.7: Part (a): Detrended averaged FHR (indirect phonocardiographic).
Part (b): SAC, SPAC, and t-statistics of SAC and SPAC of boxed segment shown in part (a); indicating an AR(1) process. The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.

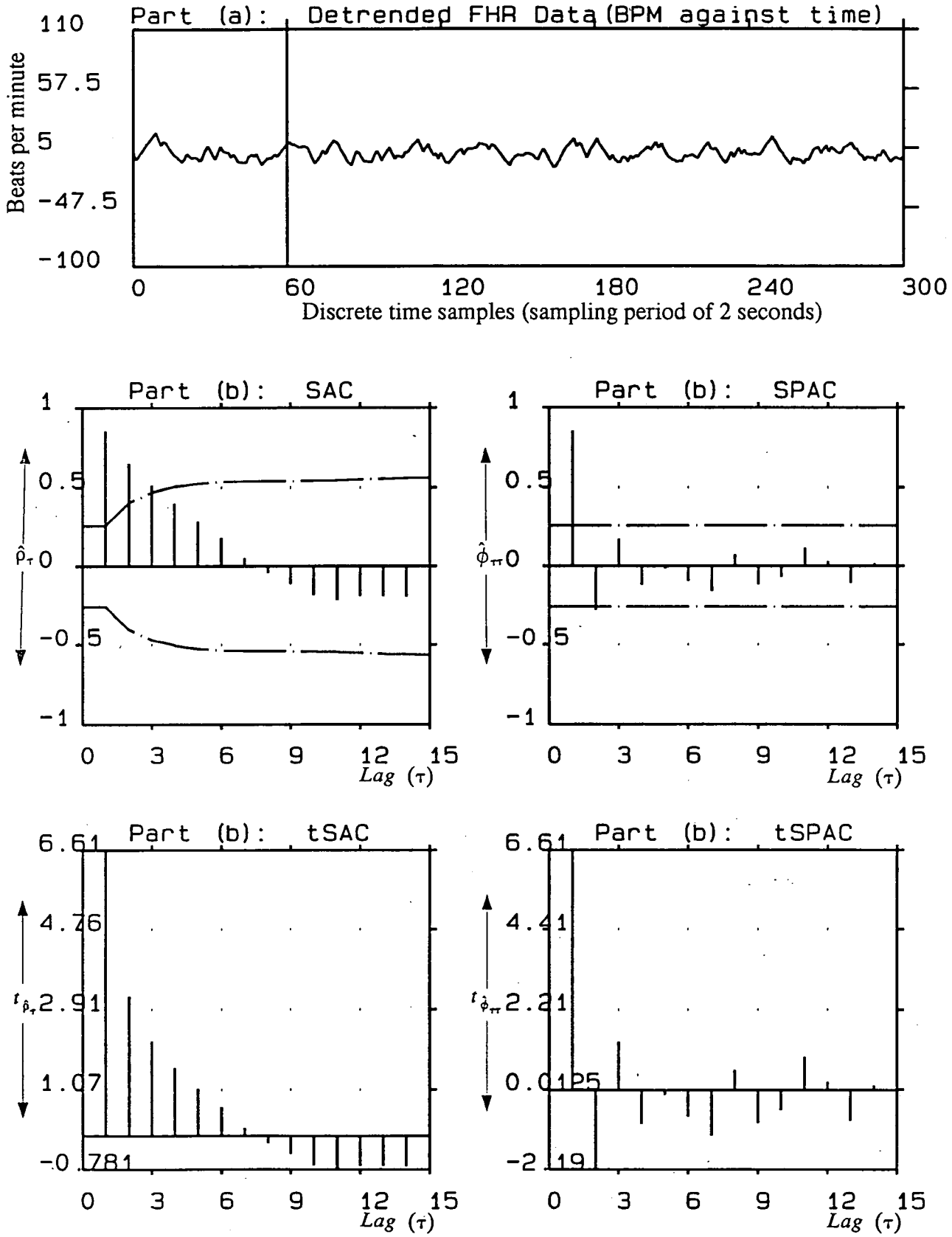


Figure 5.8: Part (a): Detrended averaged FHR (indirect phonocardiographic).
Part (b): SAC, SPAC, and t-statistics of SAC and SPAC of boxed segment shown in part (a); indicating an AR(2) process. The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.

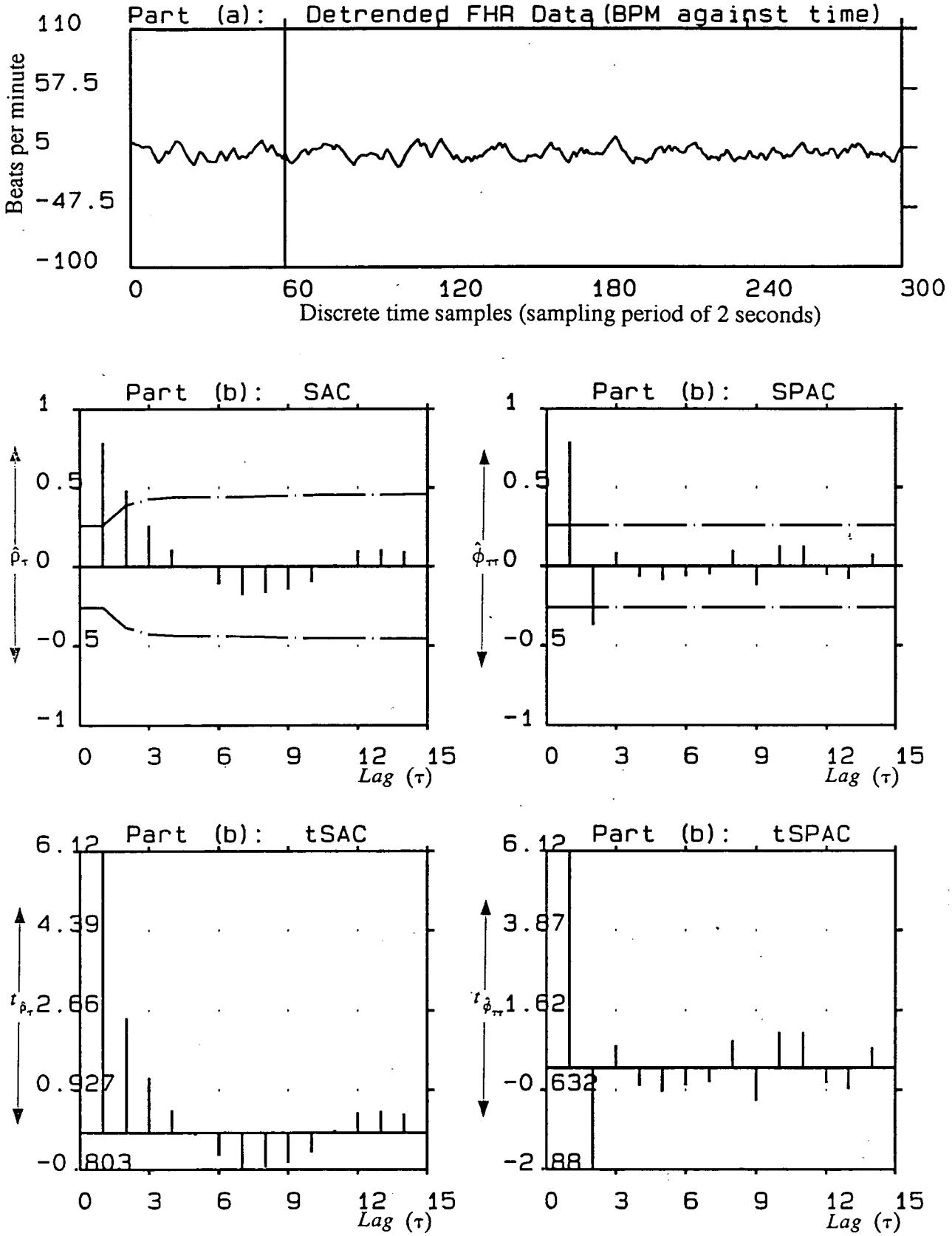


Figure 5.9: Part (a): Detrended averaged FHR (indirect phonocardiographic).
Part (b): SAC, SPAC, and t-statistics of SAC and SPAC of boxed segment shown in part (a); indicating an AR(2) process. The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.

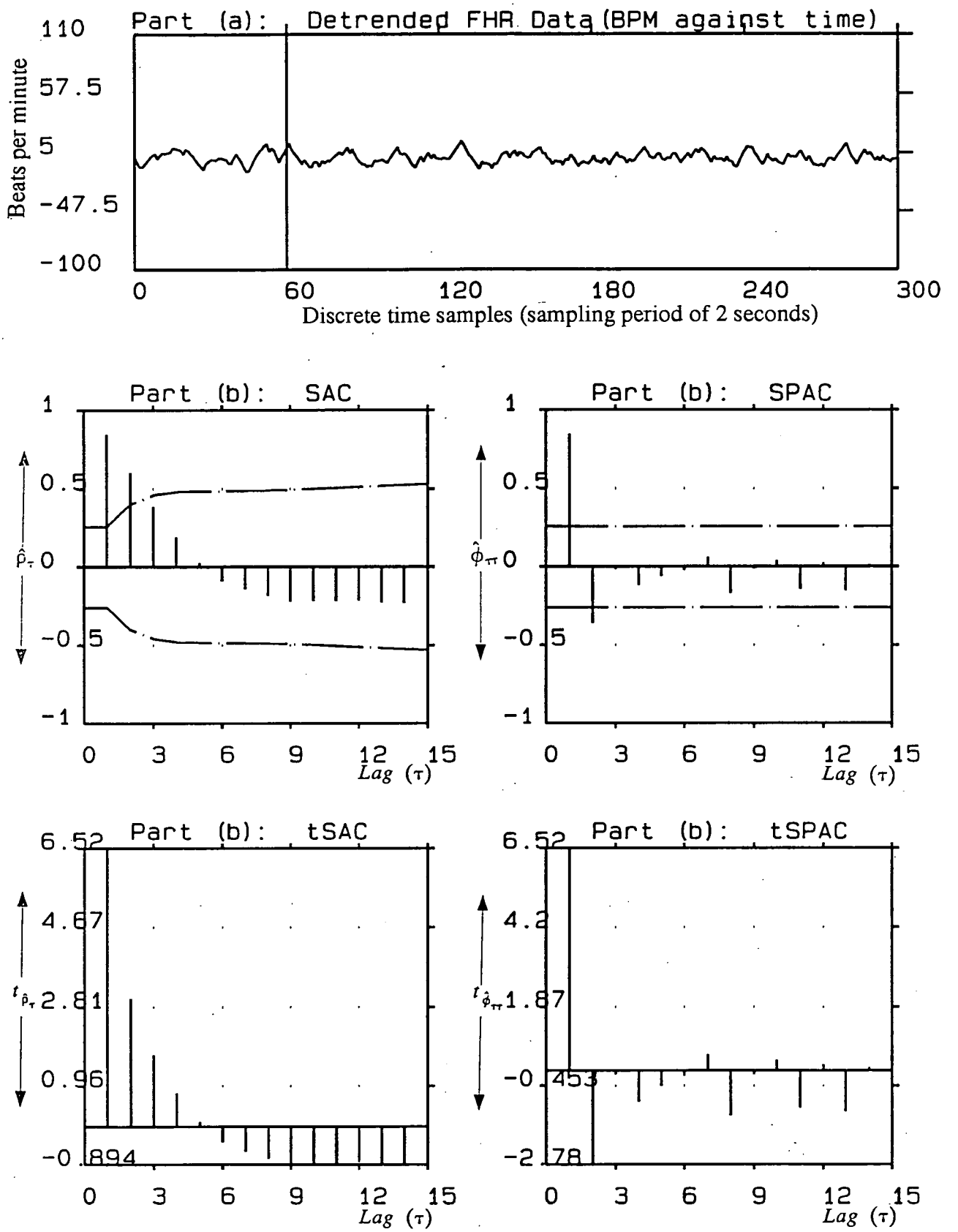


Figure 5.10: Part (a): Detrended averaged FHR (indirect phonocardiographic). Part (b): SAC, SPAC, and t-statistics of SAC and SPAC of boxed segment shown in part (a); indicating an AR(2) process. The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.

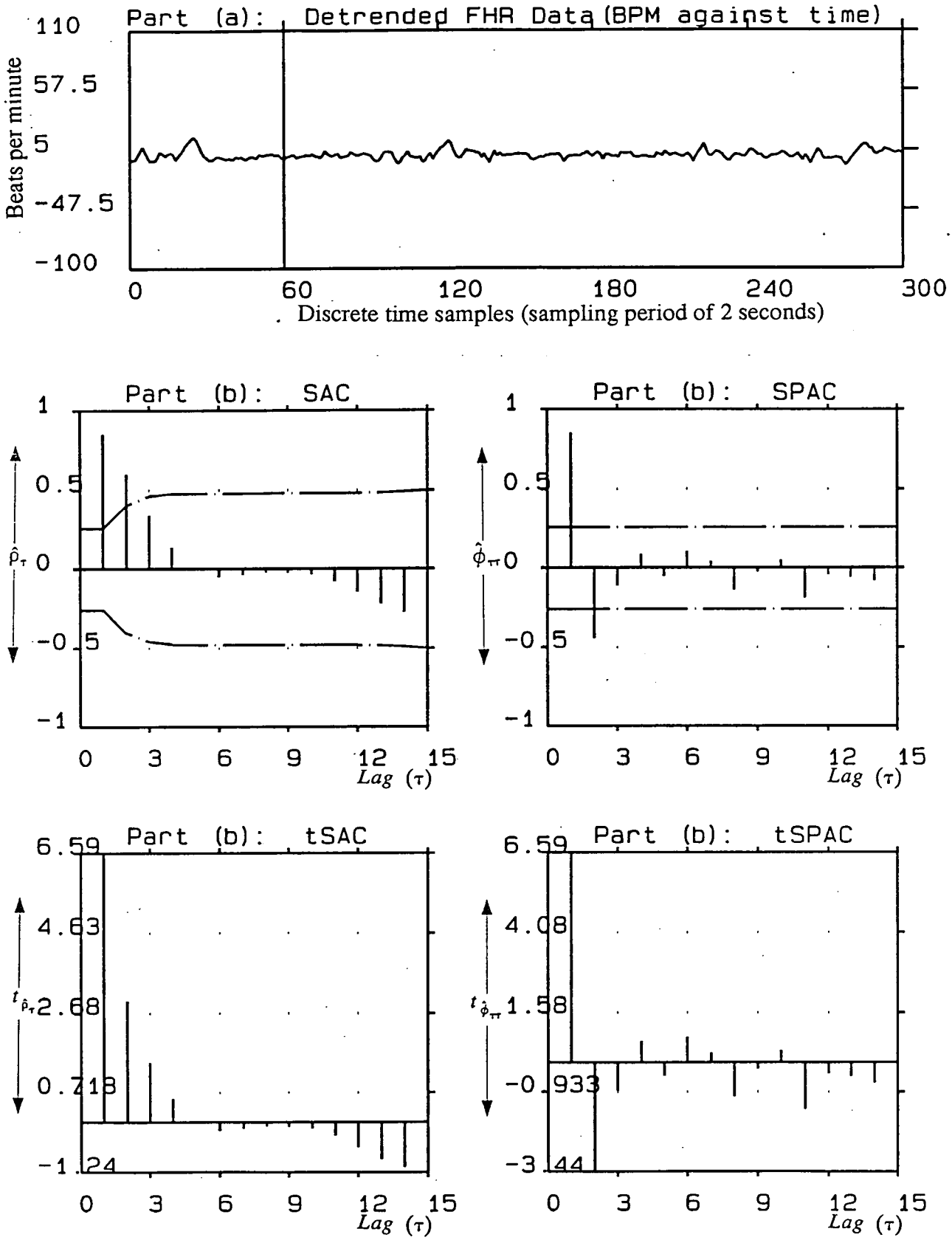


Figure 5.11: Part (a): Detrended averaged FHR (indirect phonocardiographic).
Part (b): SAC, SPAC, and t-statistics of SAC and SPAC of boxed segment shown in part (a); indicating an AR(2) process. The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.

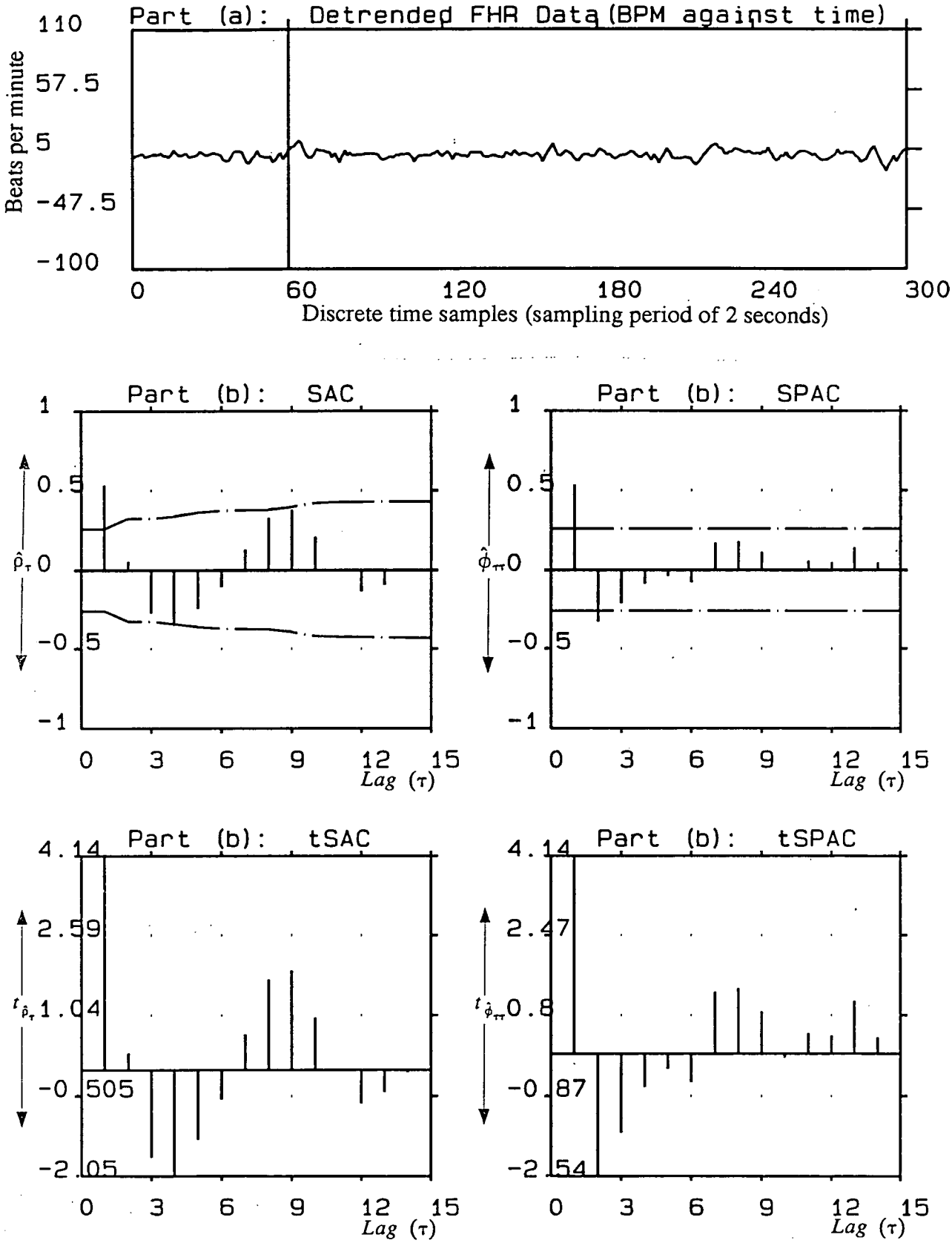


Figure 5.12: Part (a): Detrended averaged FHR (indirect phonocardiographic). Part (b): SAC, SPAC, and t-statistics of SAC and SPAC of boxed segment shown in part (a); indicating an AR(2) process. The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.

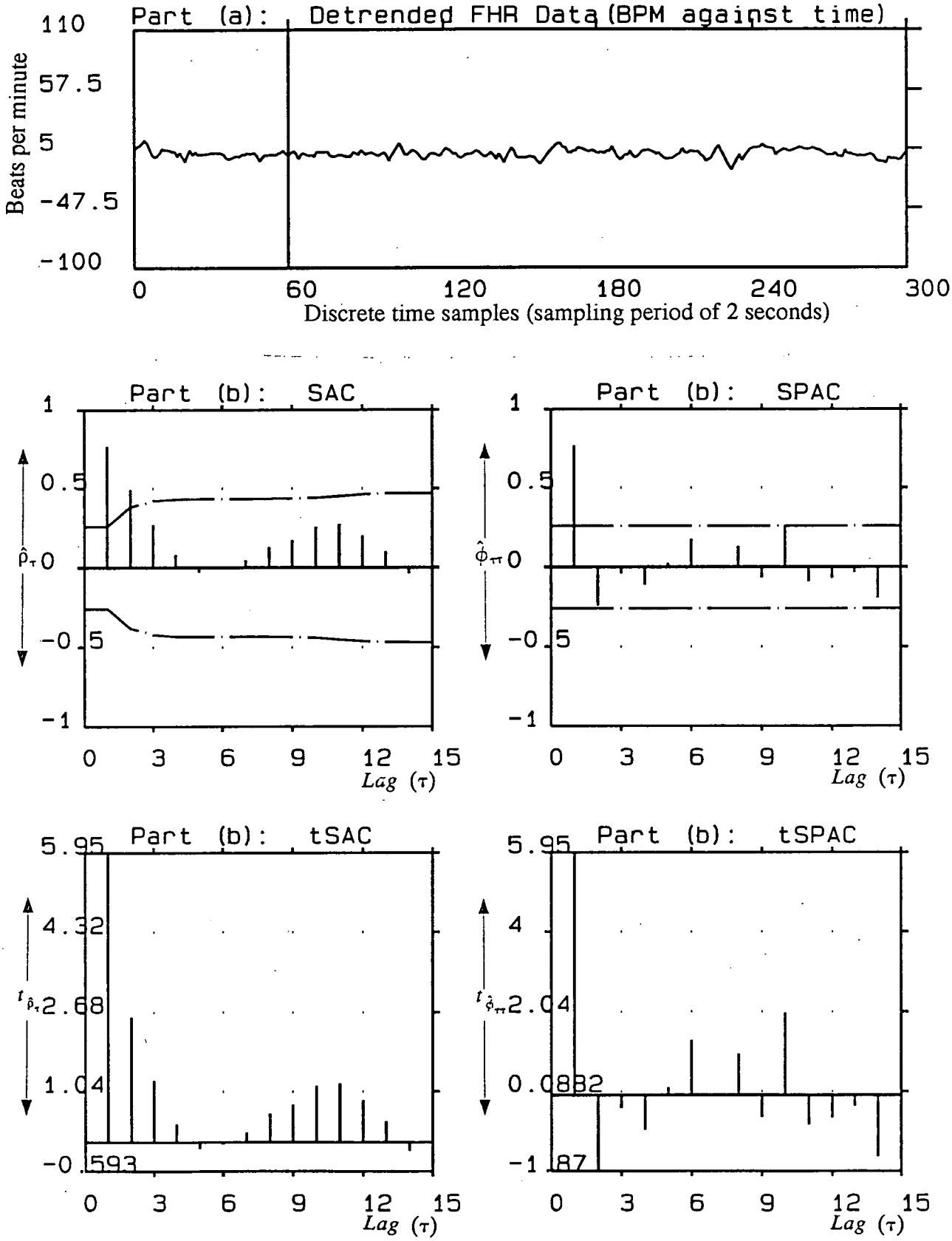


Figure 5.13: Part (a): Detrended averaged FHR (indirect phonocardiographic). Part (b): SAC, SPAC, and t-statistics of SAC and SPAC of boxed segment shown in part (a); indicating an AR(1) process. The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.

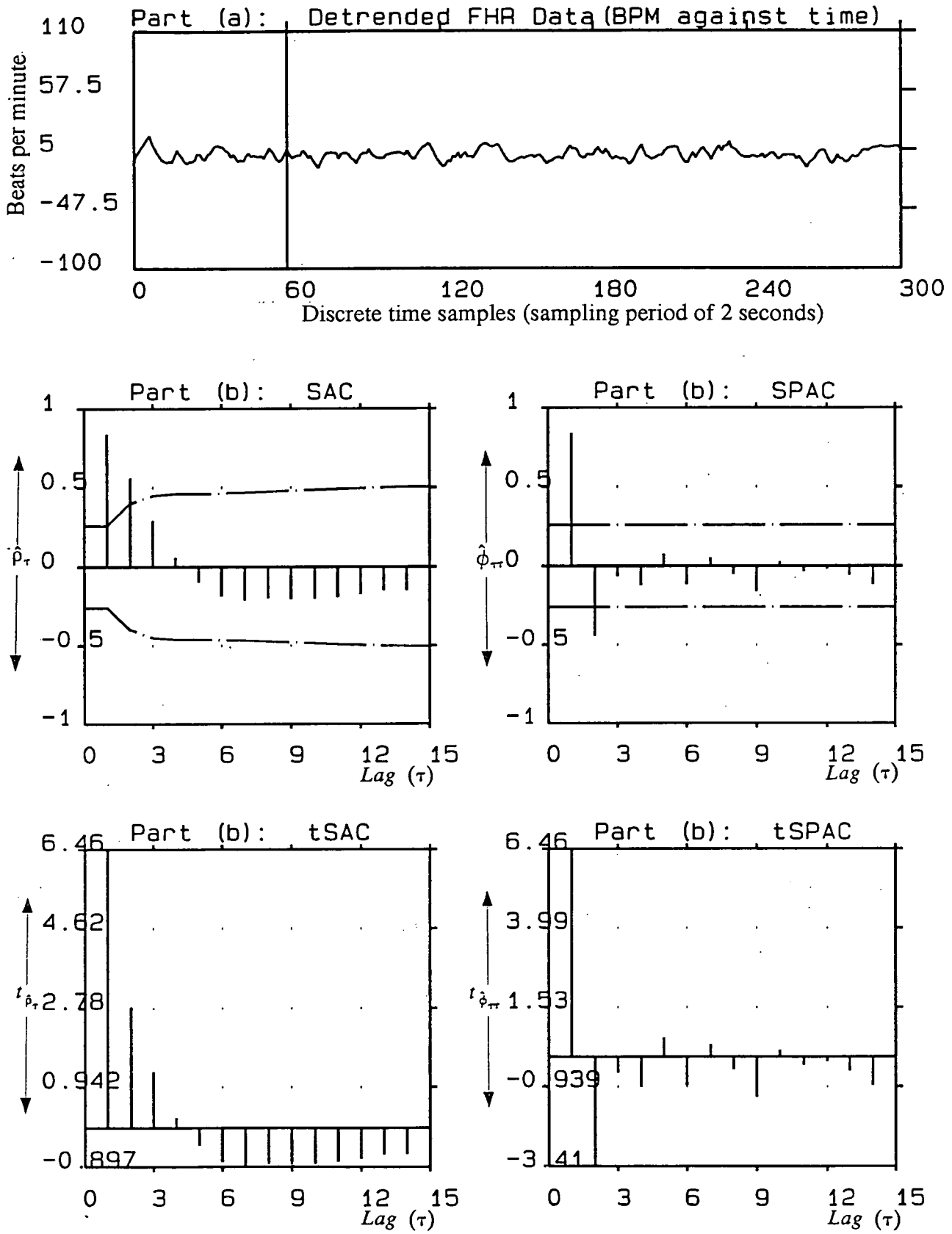


Figure 5.14: Part (a): Detrended averaged FHR (indirect phonocardiographic). Part (b): SAC, SPAC, and t-statistics of SAC and SPAC of boxed segment shown in part (a); indicating an AR(2) process. The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.

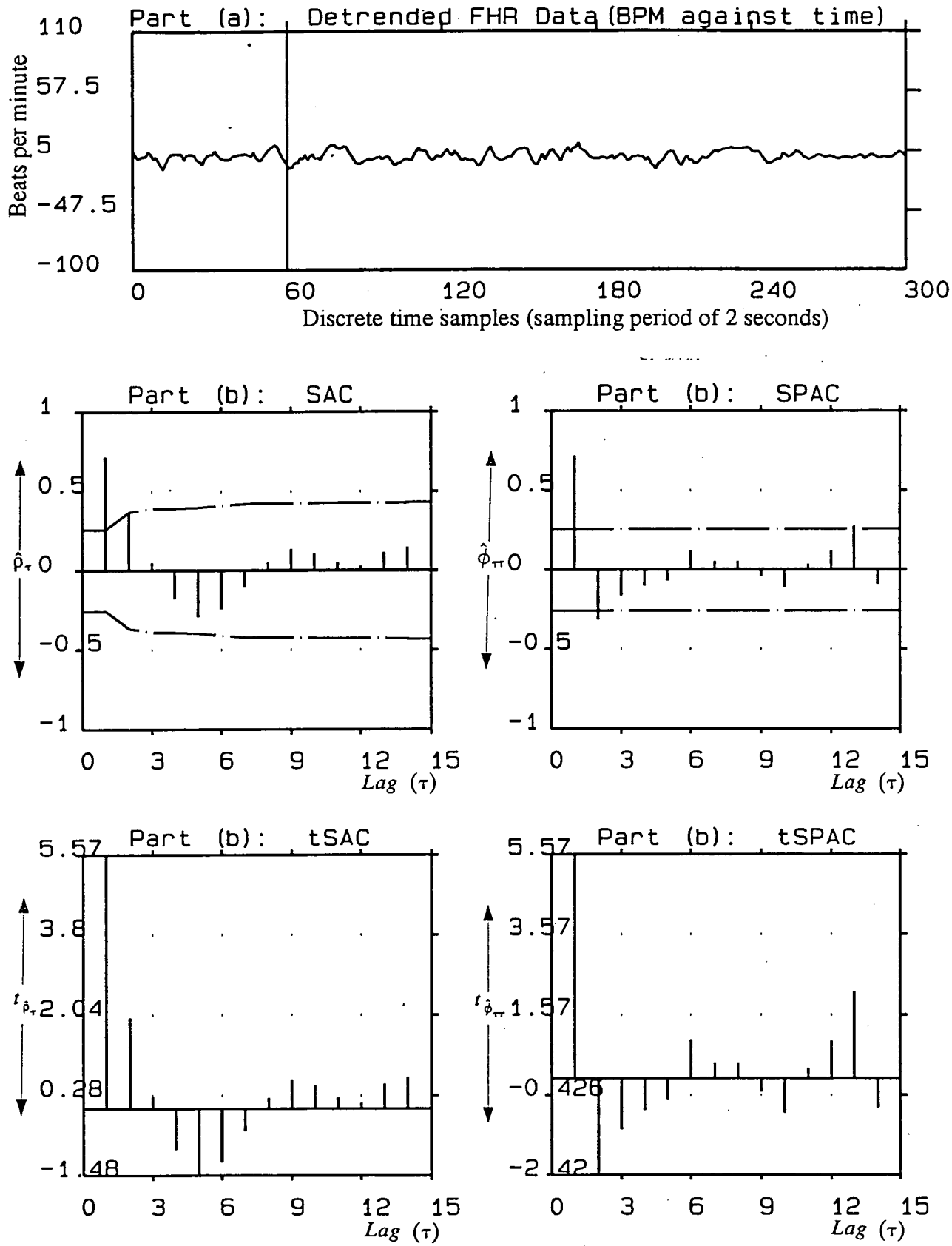


Figure 5.15: Part (a): Detrended averaged FHR (indirect phonocardiographic). Part (b): SAC, SPAC, and t-statistics of SAC and SPAC of boxed segment shown in part (a); indicating an AR(2) process. The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.

this chapter, particularly the characteristics of these processes in terms of their ACF (ρ_τ), and PACF ($\phi_{\tau\tau}$). In practice, these ACF and PACF are unknown, and for a given observation, they have to be estimated by SAC ($\hat{\rho}_\tau$) and SPAC ($\hat{\phi}_{\tau\tau}$) discussed in sections 5.3.1 and 5.4.1.

With regard to identification strategy, the approach introduced in detail in this chapter is the one which time series analysts apply in order to identify models which can parsimoniously and statistically adequately represent their data. To the naked eye, random irregular fluctuations comming from many different types of stochastic processes usually look visually similar. One cannot just guess a right parsimonious stochastic model by looking at these fluctuations.

The technique used here reveals the type of model that can represent the underlying mechanisms. Alternative identification routines do not provide clear guidance towards the type of stochastic model structure which could represent this data. They often operate via trial and error, looking blindly at a list of possible model structures (AR,MA or ARMA) and orders. For the detrended FHR data that were used in this analysis, and applying the identification criteria introduced in this chapter, the SAC and SPAC patterns that were observed unanimously supported pure AR type models.

In choosing a model, one must attempt to adhere to the principle of *parsimony*; that is, the model used should require the smallest possible number of parameters that will adequately represent the data. The objective of this type of time series analysis is not to exactly reconstruct the time series itself as in analytical methods (high resolution spectral analysis)[48, 81], but to extract just the general features which will help us to discriminate quantitatively between different patterns observed in the time series. This is the general objective of this project with regard to long term FHR variability.

A tentative model had to be identified which represented parsimoniously the correlation signature (SAC and SPAC patterns) of the majority of short, locally stationary, blocks observed. Without any doubt, the tentative model which could parsimoniously represent the underlying process was found to be a second order autoregressive structure (AR(2)). The adequacy of which was further tested at the model fitting and diagnostic stage of the time series analysis which will be introduced in the next chapter.

CHAPTER 6

MODEL FITTING, DIAGNOSTICS RESULTS AND DISCUSSION

6.1 Introduction

After a model has been tentatively specified, the next step is to estimate the parameters in the model. This stage is known as *model fitting* in time domain time series analysis. Model fitting consists of finding the best possible estimates of those unknown parameters within a given model. In this chapter a section has been allocated to introduction of the model fitting methodology that was used.

Once parameters have been estimated, *diagnostic* checks must be applied to the model to determine its adequacy. *Model diagnostics* is concerned with analyzing the quality of the model that has been identified and estimated. How well does the model fit the data? Are the assumptions of the model reasonably satisfied? If no inadequacies are found, the modeling may be assumed to be complete. The model can be used to forecast future values, or in the case of this project, the estimated model parameters can be used as numerical features representing non-white patterns exhibited by the long term variability component of FHR. Model diagnostics is another topic which will be covered in this chapter.

For the model fitting and diagnostics stages, examples will be provided to further

prove the adequacy of the tentatively identified AR(2) structure for the purpose of parsimonious parametric representation of long term FHR variability patterns. Finally the concepts of *stability* and *pseudo –periodicity* will be introduced towards the end of the chapter.

6.2 Model Fitting

There are many ways of estimating the model parameters. In time series analysis, by far the most general and powerful method of estimation is that known as *maximum likelihood* (ML)[42, 58, 60, 62, 64, 82, 83, 89-91]. This method produces almost all the well known estimates, and more importantly, may be applied (in principle) to any type of estimation problem *provided* only that the joint probability distribution of the observations is known.

The method of ML attempts to incorporate *all the information* in a model by working with the complete distribution of the observations, unlike the other inferior estimators (like method-of-moments estimators and least-squares estimators) which base their estimation on just the first two moments of the observations.

One of the main drawbacks in using ML estimation approach is its much higher level of complexity due to the involvement of joint probability density functions. This leads to criterion functions which can only be solved iteratively using nonlinear optimization techniques which in itself is a vast and complex topic.

Nevertheless, the fact that the best estimator of a parameter is non-linear should not be regarded as a deterrent to employing that estimator. Indeed, when faced with a new situation, there is a strong argument for basing estimation of the parameters of a tentatively specified model on ML, rather than attempting to derive simpler techniques which must then be shown to be asymptotically equivalent to the ML

estimators. In the next section the *likelihood* concept will be explained in greater detail as a prelude to the implementation of full ML estimation which will make use of the *Kalman filtering* technique as part of its structure.

6.3 Understanding the Likelihood Concept

The *likelihood* concept was first proposed by R.A. Fisher[42], for the purpose of estimating *deterministic parameters* that represent the distribution of statistically independent, identically distributed, random variables.

For better understanding, first consider a vector of unknown parameters Θ that describes a collection of N independent, identically distributed, observations $z(k)$, $k = 1,2,\ldots,N$. First Collecting N of these measurements into a $N \times 1$ vector $\mathbf{z}$,

$$\mathbf{z} = col[z(1),z(2),\ldots,z(N)]. \tag{6.1}$$

The *likelihood* of Θ, given the independent observations $\mathbf{z}$, is defined to be proportional to the value of the probability density function of the observations given the parameters:

$$l[\Theta|\mathbf{z}] \propto p(\mathbf{z}|\Theta), \tag{6.2}$$

Where l is the *likelihood* function representing the likelihood of Θ given the observation $\mathbf{z}$, and p is the conditional joint probability density function. On the condition that $z(k)$ are independent and identically distributed, l can be defined as,

$$l[\Theta|\mathbf{z}] \propto p(z(1)|\Theta)p(z(2)|\Theta)\ldots p(z(N)|\Theta). \tag{6.3}$$

It is important to mention here that, although the formula for the conditional joint probability density function of a realization is exactly equivalent to the formula for likelihood function, the way that the formula is interpreted is different. In the case of

the conditional joint probability density function, the parameters Θ are considered as fixed and the $z(k)$'s (representing the realization) as variable. In the case of the likelihood function, the values of the parameters can vary but the $z(k)$'s are fixed numbers as observed in a particular sample. This is the best way to visualize the structure of likelihood function.

In many situations $p(\mathbf{z}|\Theta)$ is exponential (e.g., Gaussian). It is easier then to work with the natural logarithm of $l[\Theta|\mathbf{z}]$ than with $l[\Theta|\mathbf{z}]$. Thus let

$$L[\Theta|\mathbf{z}] = \ln l[\Theta|\mathbf{z}]. \tag{6.4}$$

The quantity L is referred to as the *log −likelihood* function. The logarithm of l is a monotonic transformation of l (i.e., whenever l is decreasing or increasing, ln l is also decreasing or increasing); therefore, the point corresponding to the maximum or minimum of l is also the point corresponding to the maximum or minimum of L.

In summary the *maximum likelihood estimate* (MLE) $\hat{\Theta}$ is the value of Θ that optimizes the log-likelihood, L, for a particular set of observations $\mathbf{z}$. Practically, there are two main problems in obtaining MLE's of parameters in models of dynamical systems (ARMA,AR,..etc): (1) obtaining an expression for $L(\Theta|\mathbf{z})$ and (2) maximizing (or minimizing depending on how L is defined) $L(\Theta|\mathbf{z})$ with respect to Θ. Initially the first problem will be addressed, and afterwards the optimization approach will be considered.

6.3.1 Derivation of the Log-Likelihood Function

In a time series model, the observations are by definition dependent. Hence equation (6.3) is not applicable. The ML theorems, as presented, are based on the assumption that the observations are independently and identically distributed. This section is con-

cerned with the practical question of constructing the likelihood function for a set of dependent observations. This situation of having dependent observations is the normal case in time series analysis.

This problem of statistical dependency of a time series in the observed windows has been remedied via the use of the innovations process (or prediction errors) generated by a *Kalman* filter. As a prelude to this final goal, the dynamic time series model must first, be represented in a state space format (or Markovian representation), since the matrices that appear in the state space representation are used to construct the Kalman filter[47, 62, 92].

The standard state space representation of a dynamic model[42, 47, 62, 92], is as follows:-

$$x(k+1) = \Phi x(k) + \Gamma e(k), \tag{6.5}$$

which is the system or message model, and

$$z(k+1) = Hx(k+1) + a(k+1), \tag{6.6}$$

is the observation equation or model. The initial state vector $x(0)$ is Gaussian, and has an $(n \times 1)$ dimension, e is the input noise vector $(p \times 1)$, z is the observation vector $(m \times 1)$, a is the observation noise vector $(m \times 1)$, Φ is the system matrix $(n \times n)$, Γ is the input noise matrix $(n \times p)$, H is the observation matrix $(m \times n)$ and the vector terms in $e(k)$ and $a(k)$ are mutually stationary Gaussian white noise sequences, for which $E[e(k)e'(k)] = Q$ and $E[a(k)a'(k)] = R$, where k = 0,1,...,N-1, and

$$\Theta = [\Phi, \Gamma, H, Q, R]',$$

$$\mathbf{z} = [z(1), z(2), \ldots . z(N)]'.$$

We must obtain the joint density function $p(\mathbf{z}|\Theta)=p(z(1),z(2),...z(N)|\Theta)$. Referring to equations (6.5) and (6.6), it is clear that the N measurement (or observation) vectors $[z(1),z(2),...z(N)]$ are not statistically independent; hence,

$$p(z(1),z(2),...z(N)|\Theta) \neq p(z(1)|\Theta)p(z(2)|\Theta)...p(z(N)|\Theta).$$

This problem of statistical dependency can be solved via the *innovations process*, $\nu(k|k-1)$ of a linear Kalman filter (refer to table 6.1)[92, 93]. Innovations sequence also known as residuals or prediction error terms are in general equivalent to the difference between an original observed measurements $z(k)$ and its corresponding predictions $z(k|k-1)$, which are computed optimally in a least mean square sense via a linear Kalman filter. The word innovation denotes "newness", and this quality is represented by the fact that, the optimally estimated process $\nu(k|k-1)$, will be white and statistically independent in characteristics, and any redundant information in the form of correlation in the measurements $z(k)$ has been removed. Hence only new information is retained in the innovations process.

In general the innovations process $\nu(k|k-1)$ and measurements $z(k)$ are probabilistically equivalent, and as a result are said to be causally invertible[90, 94, 95], that is $\nu(k|k-1)$ is computable from $z(k)$, and $z(k)$ is computable from $\nu(k|k-1)$; thus, the conditional probability density function, $p(\nu(1|0),\nu(2|1),...,\nu(N|N-1)|\Theta)$, contains the same data information as $p(z(1),z(2),...,z(N)|\Theta)$ does. As a result $L(\Theta|\mathbf{z})$ can be represented as:

$$L(\Theta|\mathbf{z}) = L(\Theta|\nu) = \ln p(\nu(1|0),\nu(2|1),...,\nu(N|N-1)|\Theta).$$

Now for an optimal filter, the innovations is a zero-mean Gaussian white noise sequence, so that the log-likelihood can be represented[58, 62, 90, 92] as,

$$L(\Theta|z) = -\frac{N}{2}ln(2\pi) - \frac{1}{2}\sum_{k=1}^{N}[\nu(k|k-1)V_{\nu}^{-1}(k|k-1)\nu(k|k-1) + ln|V_{\nu}(k|k-1)|], \quad (6.7)$$

where $V_{\nu}(k|k-1)$ is the innovations process covariance. The constant term $-(N/2)ln(2\pi)$ can be omitted, because it does not depend on Θ.

As previously stated, the innovations sequence and its covariance can be generated by a *Kalman filter* which can be implemented from the dynamics and statistics that represent our model. Table 6.1 shows implementation of a *linear* Kalman filter structure, for an AR(2) + observation noise (σ_a^2) time series model. Because the state space representation also usually allows for measurement error or observation noise, we do not observe $x(k)$ but only observe $z(k)$ through the observational equation (6.6), where $H = [1\ 0]$ for an AR(2) model and $a(k)$ is another zero mean, white noise process which is independent of $e(k)$.

In the case of time series analysis sometimes there is no observational noise in the time series. In this case there is no need to model observational noise. This applies to the AR(2) structure identified in the previous chapter for long term FHR variability. Hence, the observation noise term $a(k)$ is set to zero.

This is reflected in the linear Kalman filtering equations by setting the variance, σ_a^2, of the observation noise or (measurement error) to zero[58, 62, 96]. This linear structure is then used to provide information needed for implementation of the log-likelihood criterion function shown in equation (6.7).

In our situation the true Θ_T is not known, but needs to be estimated from the maximization of $L(\Theta|z)$. At the start, the Kalman filter will behave suboptimally, since wrong values of Θ are being used by it, but as better estimates $\hat{\Theta}$ are given to the Kalman filter, the estimates of state vector $x(k)$, improve. In fact, if $\hat{\Theta} \to \Theta_T$ as $N \to \infty$, then

Prediction:-

State Prediction $\hat{x}(k|k-1) = \Phi\, \hat{x}(k-1|k-1)$

Prediction error covariance $V_{\tilde{x}}(k|k-1) = \Phi\, V_{\tilde{x}}(k|k)\, \Phi^T + \Gamma\, Q\, \Gamma^T$

where Φ is the Signal Process Matrix, Γ is Input Matrix and Q is Input Noise variance Matrix as shown below.

$$Q = \begin{bmatrix} \sigma_e^2 & 0 \\ 0 & 0 \end{bmatrix} \qquad \Phi = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix} \qquad \Gamma = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Update:-

Kalman Gain $K(k) = V_{\tilde{x}}(k|k-1)\, H^T\, V_\nu^{-1}(k|k-1)$

Innovation Covariance $V_\nu(k|k-1) = HV_{\tilde{x}}(k|k-1)H^T + R$

Estimation error covariance $V_{\tilde{x}}(k|k) = V_{\tilde{x}}(k|k-1) - K(k)HV_{\tilde{x}}(k|k-1)$

Estimation or Correction $\hat{x}(k|k) = \hat{x}(k|k-1) + K(k)\nu(k|k-1)$

Innovation or Prediction error $\nu(k|k-1) = z(k) - \hat{z}(k|k-1) = z(k) - H\hat{x}(k|k-1)$

where $H = [\,1 \;\; 0\,]$ is the observation matrix.

The Kalman Filter is completely specified by four parameters:

$$\Theta^T = [\phi_1, \phi_2, \sigma_e^2, \sigma_a^2]$$

The initial values, for $\hat{x}(0|0)$ and $V_{\tilde{x}}(0|0)$ in this algorithm are also necessary. For a stationary realization, the theoretical mean value and the covariance matrix of the state vector are used. These are computed from the identified model[62, 96].

Table 6.1 Kalman Filtering Equations

$\nu_{\hat{\theta}}(k|k-1) \rightarrow \nu_{\theta_T}(k|k-1)$ as $N \rightarrow \infty$, Then the sub-optimal Kalman filter will approach the optimal Kalman filter.

Finally, this form of ML parametric estimation is adaptable to situations where noisy data might arise. This is achieved by allowing an observation noise term to be used as part of the time series model for the purpose of parsimonious representation. Also the use of a Kalman filter as part of the parametric estimator allows the algorithm to cope with measured time series which have missing data [62, 96, 97]. Missing data occasionally occurs in FHR time series, particularly on rapid FHR decelerations as the correlation based FHR estimator produces a gross error (outlier) which is recognized as such by post processing circuitry and suppressed by the machine producing no output. In terms of hard copy (plotter) output this is achieved by lifting the plotting pen.

6.4 Optimization Method

There is no closed-form solution for Θ that maximizes $L(\Theta|\mathbf{z})$. The maximum likelihood estimate (MLE) $\hat{\Theta}$ is a non-linear optimization problem.

The objective is to change $\hat{\Theta}$ iteratively, so as to optimize $L(\Theta|\mathbf{z})$ subject to the constraints of a Kalman filter. This optimization is achieved via the Kalman filter, which generates $V_\nu(k|k-1)$ and $\nu(k|k-1)$, both being dependent on Θ, and enabling calculation of $L(\Theta|\mathbf{z})$.

The algorithm applied to optimize the likelihood function, was a *non −gradient* technique. The method used was Powell's direction set method in multidimensions. This is the most widely accepted direct search method. It is a powerful method and has been proven to be superior to some of the descent methods[55, 98]. Only function values are used, and minimization is achieved along an artfully chosen set of favorable directions. Basically for each direction found, the algorithm moves along that direc-

tion to its minimum. Then from there along the second direction to its minimum, and so on, until the function stop decreasing.

There are two different approaches in updating the search directions, both of which have been tackled by Powell. The first is based on the concept, of *non −interfering* directions, more conventionally called *conjugate directions* [55, 98-100]. Theoretically if one performs successive line minimizations of a function along a conjugate set of directions, then one doesn't need to redo any of those directions. The other way is more *heuristic* [55, 98, 100, 101]. It tries to find a few good directions far along narrow valleys. This later scheme has been used to optimise the log-likelihood function, since it tries to avoid building up a set of directions which fold up on each other and become linearly dependent, resulting in failure to find the global minimum, and will be explained in the next section.

6.4.1 Powell's Direction Search Routine Applied

In this section the actual algorithm that was applied will be explained in greater detail. The procedure is a *direction search method*, and does not require any derivative computations. The Powell's algorithm that defined the applied minimization subroutine proceeds as follows:

1). A *starting point*, $\Theta_0^{(0)}$ is selected. The starting point that was used and given to the minimization routine is based upon the moment based estimates, derived from the Levinson-Durbin recursion introduced earlier. This considerably reduces the minimization iteration number required for the algorithm to converge. This form of moment based setting of the starting point of the minimization is especially valuable and significant when the AR model order is high. The initial search directions from this starting point, $S_i^{(0)}$, $i = 1,2,\ldots,P$, are parallel to the original coordinate axes (where P is

number of parameters that must be estimated).

2). A sequence of single variable searches are made in the P initial directions. This has been achieved by bracketing the minimum in each of these directions, and using *inverse parabolic interpolation* to zoom in on the minimum[55, 100].

3). The following points are then located:-

$\Theta_P^{(k)}$ = The last point from the sequence of single variable searches,

$\Theta_M^{(k)}$ = The point of greatest function improvement between successive single variable searches,

$\Theta_E^{(k)} = 2\Theta_P^{(k)} - \Theta_0^{(k)}$ = Expanded point,

$\Theta_0^{(k)}$ = Starting point for the iteration,

where k is the iteration stage index which is incremented for each new set of search directions.

4). A check is made to see if the value of the log-likelihood objective function at the expanded point, $\Theta_E^{(k)}$, is better than that at the starting point, $\Theta_0^{(k)}$. If there is no improvement, the last point, $\Theta_P^{(k)}$, becomes the new starting point and a new sequence of single variable searches are made in the same directions as before,

$$\Theta_0^{(k+1)} = \Theta_P^{(k)},$$
$$S_i^{(k+1)} = S_i^{(k)}, \quad i = 1,2,3,\ldots,P,$$

If the loglikelihood objective function, $L_E^{(k)}$, at the expanded point is an improvement over the starting point function, $L_0^{(k)}$, then the following test is performed,

$$[L_0^{(k)} - 2L_P^{(k)} + L_E^{(k)}][L_0^{(k)} - L_P^{(k)} - \delta]^2 \geq \delta \frac{(L_0^{(k)} - L_E^{(k)})^2}{2},$$

where $\delta = |L_M{}^{(k)} - L_{M-1}{}^{(k)}|$.

The test determines whether the function in this region is a valley, i.e, $L_E{}^{(k)}$ is an improvement but the surface is rising. If the test is satisfied, the old search directions are retained and a new sequence of single variable searches starts as above. If the test is not satisfied, a single variable search is performed in the $\mu^{(k)}$ direction,

$$\mu^{(k)} = \Theta_P{}^{(k)} - \Theta_0{}^{(k)},$$

until the best value $\Theta_0{}^{(k+1)}$ is found. New search directions are then chosen as follows:

$$S_i{}^{(k+1)} = S_i{}^{(k)}, \quad \text{for } i = 1,2,3,\ldots,M-1,$$
$$S_i{}^{(k+1)} = S_{i+1}{}^{(k)}, \quad \text{for } i = M,\ldots,N-1,$$
$$S_P{}^{(k+1)} = \mu^{(k)}.$$

A new sequence of single variable searches is then started. What steps explained so far are repeated until convergence.

5). *Convergence* is assumed and the routine is *terminated*, when the objective function based *relative error* is less than a very small scalar positive tolerance ϵ,

$$\frac{2^* |L_0{}^{(k)} - L_P{}^{(k)}|}{|L_0{}^{(k)}| + |L_P{}^{(k)}|} < \epsilon.$$

On a machine with an S decimal digit mantissa, an appropriate value of the functional tolerance ϵ would be approximately $10^{-S/2}$, and 10^{-4} otherwise[55, 102]. On this basis the functional tolerance (ϵ) was set at 10^{-8}. With this tolerance setting, after each successful termination (between 2 to 5 iterations), resulted in the final computed objective function value to have an accuracy in significant digits of approximately 10^{-8}, which is more than adequate. A flow chart illustrating the above explained procedure is shown in table 6.2.

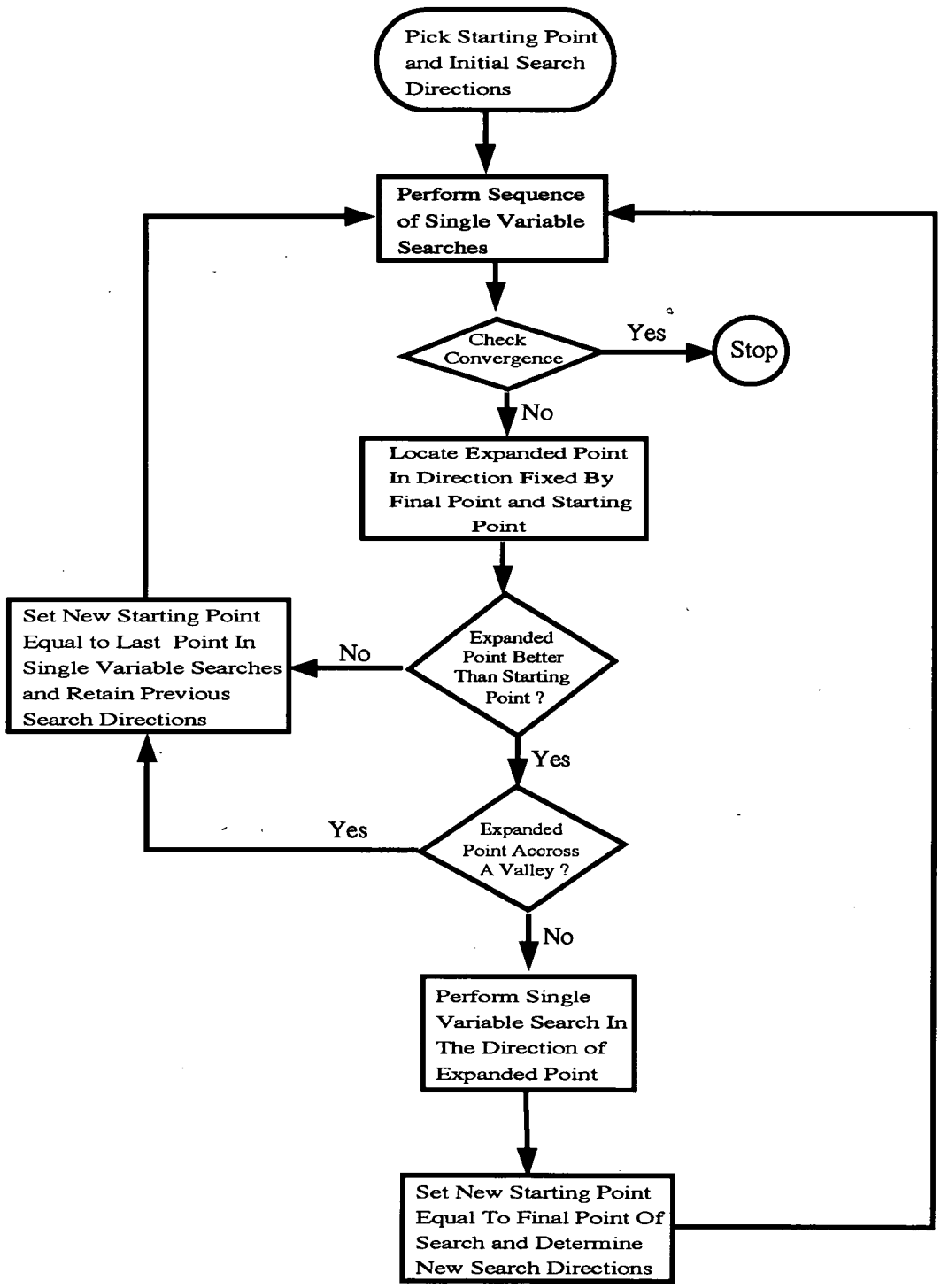


Table 6.2. General Flow Chart for the Powell's Optimization Routine .

6.4.2 Reparameterizing for Stability

One problem that was observed in optimizing the log-likelihood criterion for an autoregressive structure, was occasional parametric estimates associated with the unstable regions of the parametric space. In this situation the model will behave in an explosive manner (i.e. gives diverging mean nonstationary simulation). In order to prevent such eventualities, one has to seek some form of reparameterization of the dynamic structure so that it will prevent the nonlinear optimization algorithm from searching in regions of the parametric space associated to unstable nonstationary behaviour, resulting in meaningless parametric values.

One general condition for the stability of a pure AR structure is for the modulus of its p-th parameter, $|\phi_p|$, not to exceed 1. This follows from the fact that the p-th parameter of the AR structure is equivalent to the partial autocorrelation of a realization at the p-th lag, a quantity which never goes less than -1 or higher than +1. This is a key point, which is utilized as a step towards full reparametrization of all of the dynamic parameters of the AR structure.

In the maximum likelihood estimation of the parameters of the AR time series models, with or without observational error, the stability criterion of the equation defining the AR models can be guaranteed. This is achieved by reparameterizing in terms of the partial autocorrelation coefficients, $\phi_{\tau\tau}$, where the ϕ's are constrained to be in the open interval $(-1,1)$ by the transformation[96]:

$$\phi_{\tau\tau} = \frac{[1 - e^{-u_\tau}]}{[1 + e^{-u_\tau}]}, \quad where \ -\infty < u_\tau < \infty, \tag{6.8}$$

for $\tau = 1,\ldots,p$. The ϕ's of the AR model can then be reparameterized or transformed by the well-known Levinson-Durbin recursion for $\tau > 1$,

$$\phi_{\tau j} = \phi_{\tau-1,j} - \phi_{\tau\tau}\phi_{\tau-1,\tau-j}, \quad \text{for } j = 1,2,\ldots,\tau-1. \tag{6.9}$$

The statistical parameters, such as input noise variance σ_e^2, should be optimized in terms of their standard deviations (σ_e), in order to make sure the final estimated variance values are positive.

Finally, non-linear, unconstrained optimization can then be carried out with respect to u_τ for $\tau = 1,\ldots,p$ and σ_e for an AR(p) model. After the optimization is complete the AR(p) estimated parameters can be calculated by (6.8) and (6.9).

6.5 Diagnostic Checking

Once a precise estimation of the coefficients of the identified model has been made, diagnostic checks must be applied to the model to determine its adequacy. A statistically adequate model is one whose random shocks are independent (not correlated). In practice random shocks ($e(k)$) cannot be observed, but estimates of the random shocks are available in the form of the residuals of the estimation process (or its innovations). At the diagnostic checking stage these residuals ($\nu(k|k-1)$) are used to test hypotheses about the independence of the random shocks.

It is important that the independence of the estimated residuals be satisfied for a very practical reason. The residuals are estimates of random shocks (new independent information). These constitute the stochastic component of **z**, the realization that is being modeled. Thus, if they are serially correlated, then there is a correlation pattern in **z** that has not been accounted for by the tentatively identified model. Yet the whole idea in stochastic time series analysis is to find the right model which would avoid this eventuality.

In order to check the prediction errors (or residuals) for significant correlation, the

residuals SAC (RSAC) and the residuals SPAC (RSPAC) must be calculated together with their corresponding t-values. This is achieved using the equations (5.13),(5.23),(5.29) and (5.30) introduced in the previous chapter where they were used for model identification.

If the model is correct, then there should not be any significant correlation spikes from the RSAC and RSPAC. The computed spikes should be distributed, approximately normally about zero. They should lie within their ± 2 standard deviations ($\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ respectively; refer to (5.17) and (5.24)). Also their computed t-statistic values should all be less than two[60, 62, 82].

A more formal test, which is the main diagnostic tool used by the time series analyst is the *Portmanteau lack of fit test*. This is also known as *Box −Pierce Q −test*. The test is based upon the computation of a Q-statistic which takes into consideration the K computed RSAC lag values as a group.

The significance and accuracy of the test was further improved by Ljung and Box[58, 62, 63, 103, 104], thus the statistic is now known as *Ljung −Box* statistic and this updated version is defined by

$$Q = N\,(N+2)\sum_{\tau=1}^{K}\left(\frac{1}{N-\tau}\,\hat{\rho}_\tau^{\,2}\right), \tag{6.10}$$

where $K = N/4$ and $\hat{\rho}_\tau$ is RSAC at the lag τ. Under the null hypothesis that the model is correctly specified, Q is asymptotically distributed as a chi-square (χ^2) random variable with $(K-p-q)$ degrees of freedom.

Fitting an erroneous model causes the residuals to be more related, thus Q inflates. In general a model can be rejected if its Q statistic is greater than the *chi −square critical value* $\chi^2_{0.05}[K-p-q]$.

6.6 Applications to Real Data

In this section a series of 15 examples have been presented (self-explanatory figures 6.1 to 6.15). These are related one-to-one with the examples presented in the previous chapter. For each two minute data block fitted to the identified second order structure, the type of simulation that the AR(2) can generate from the estimated parameters has also been presented. This may be used as further visual verification of the adequacy of AR(2) structure as the simulated sequences represent the general salient features observed in the windows of observation on the real data. Along with each simulation, the residuals left after each fit with their RSAC, RSPAC, t-statistics of RSAC and RSPAC, and Q-statistics have also been provided.

Following inspection of the examples shown in figures 6.1 to 6.15, the adequacy of a second order autoregressive structure (AR(2)) for modeling long term variability cannot be rejected. Part (b) of these figures show synthetic data generated from an AR(2) model. These are produced from the estimated parameters (quoted in each of the figures) extracted from the boxed segments of detrended averaged FHR (phonocardiographic) shown in part (a) of each of the figures. The visual resemblance between the variability in the real data and its corresponding synthetic version is apparent.

Furthermore part (e) of each of the figures show the SAC, tSAC, SPAC and tSPAC of the model residuals (part (d)). All showed no significant correlation spikes, indicating the white characteristics of the estimated residuals, thus showing adequate fit of the model to data. Also the overall Portmanteau lack of fit test (or Ljung-Box test) showed positive results for all of these examples, by having their Q-statistics (as presented in part (e)) well below the chi-square critical value for 13 degrees of freedom ($\chi^2_{0.05}[13] = 22.3621$) which is an indicator of the adequacy of fit and thus the model.

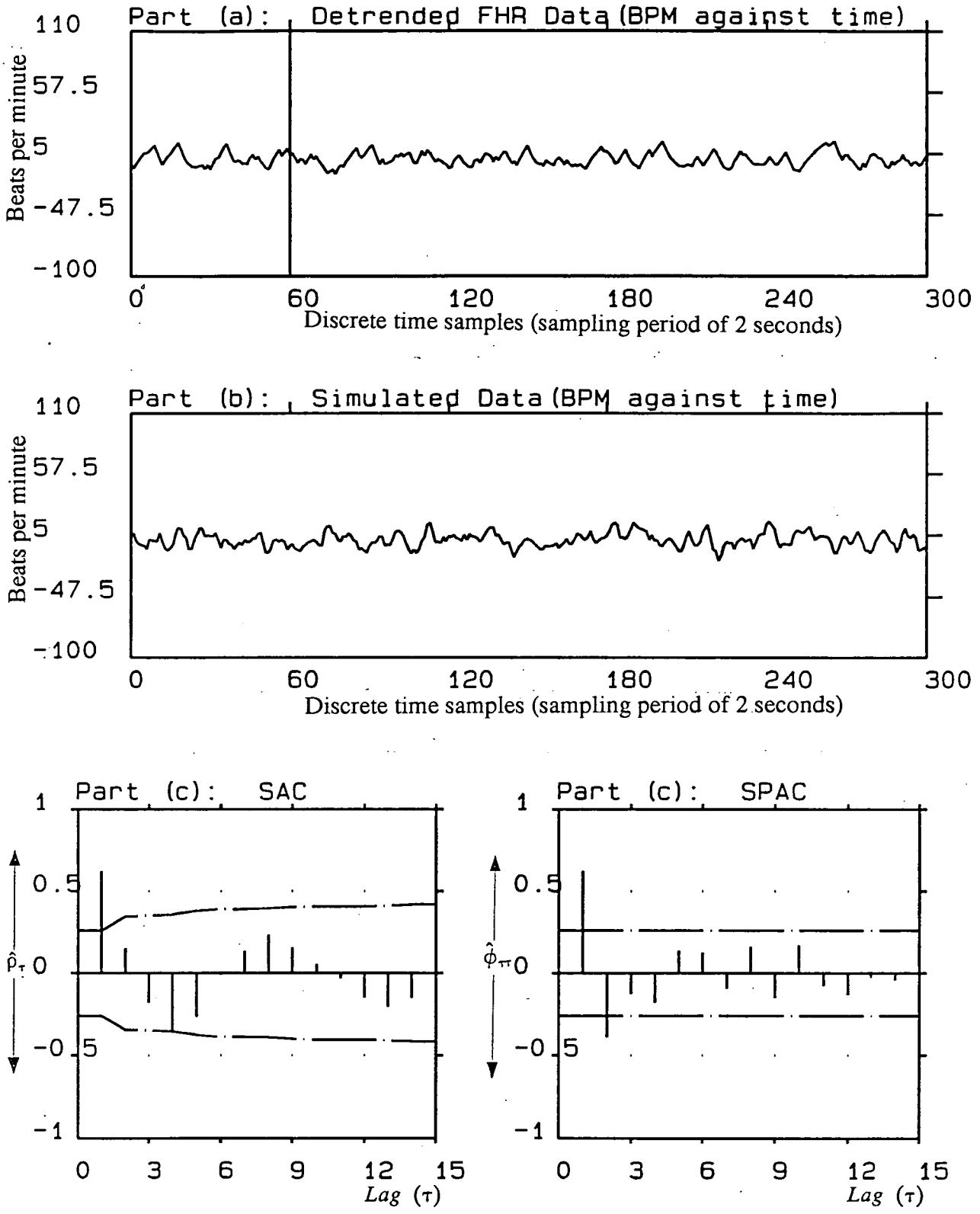


Figure 6.1: Part (a): Detrended averaged FHR (indirect phonocardiographic);
Part (b): Synthetic FHR data, generated from the parameters extracted from the boxed segment shown in part (a); the estimated parameter values were: $\hat{\phi}_1 = 1.0891$, $\hat{\phi}_2 = -0.3900$, $\hat{\sigma}_e = 3.4518$, $\hat{\sigma}_p = 6.0331$, $\hat{\mu}_{baseline} = 126.6921$; *Pseudo –periodicity* was detected ($t_0 = 51.60$ *sec*).
Part (c): SAC and SPAC of the first 60 synthetic data samples shown in part (b), indicating AR(2) correlation characteristics; The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.

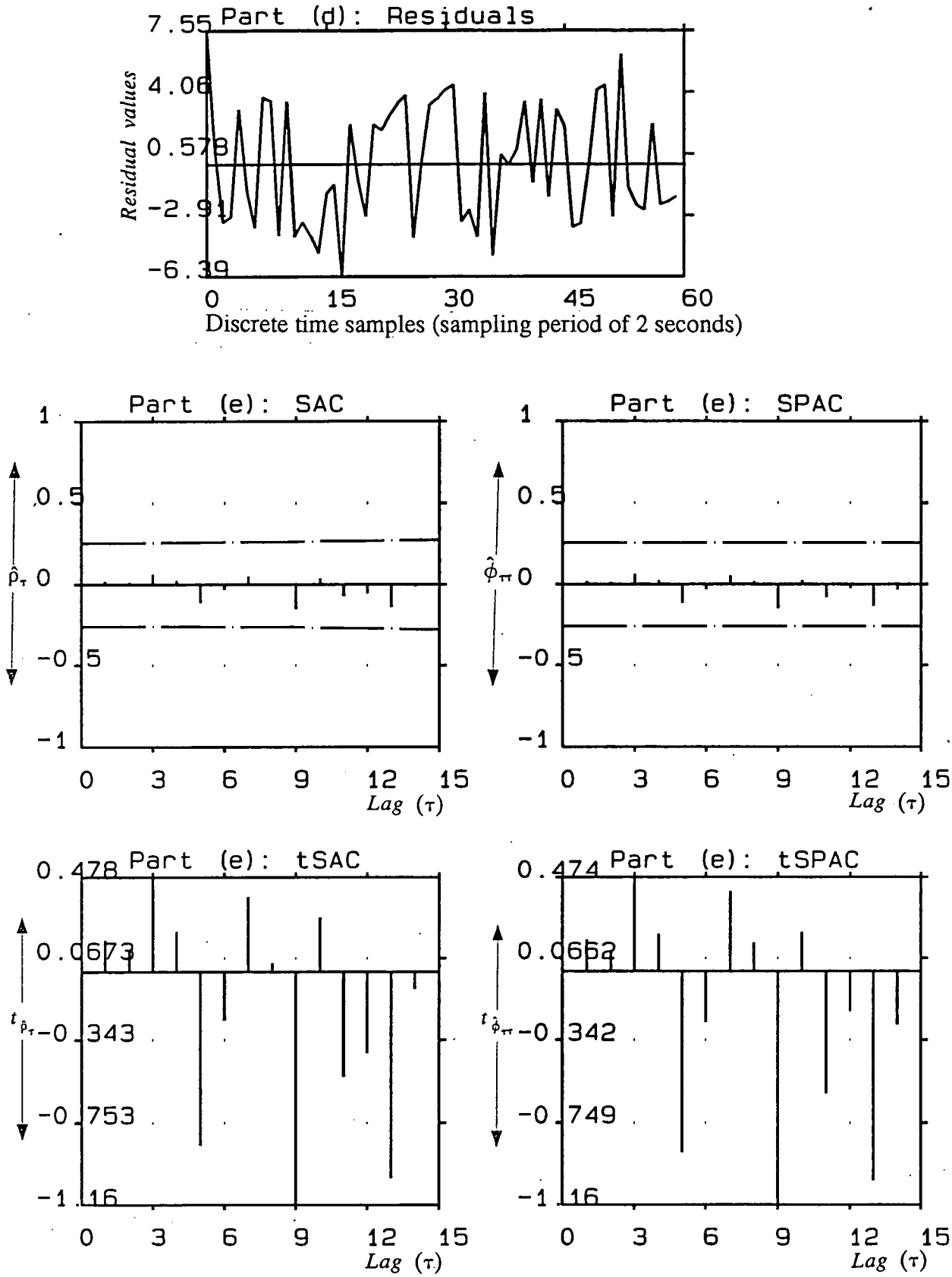
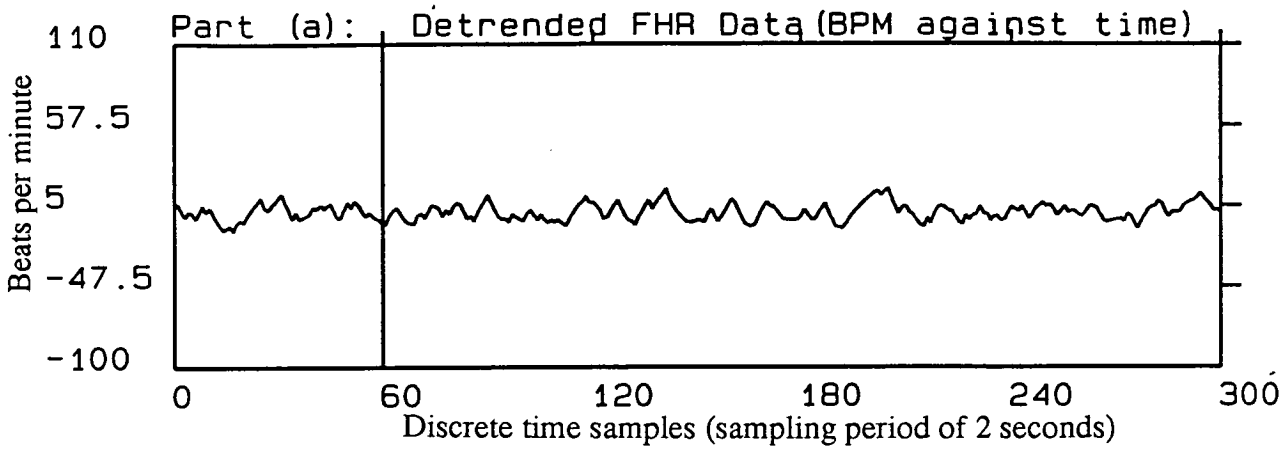
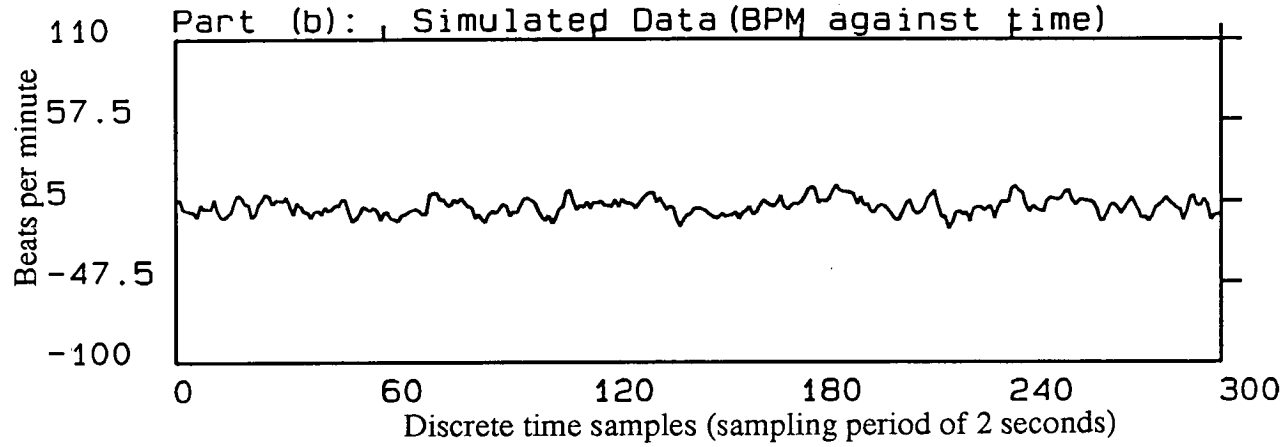


Figure 6.1: Part (d): Residuals (or Prediction errors or innovations) left after fitting an AR(2) model to the boxed segment shown in part (a).
Part (e): SAC, SPAC, t-statistics of SAC and SPAC of residuals shown above, indicating a white sequence; Q *–statistic* = 10.1165 $<\chi^2_{0.05}[13]=22.362121$, further proof of the adequacy of the model fitted to the data.





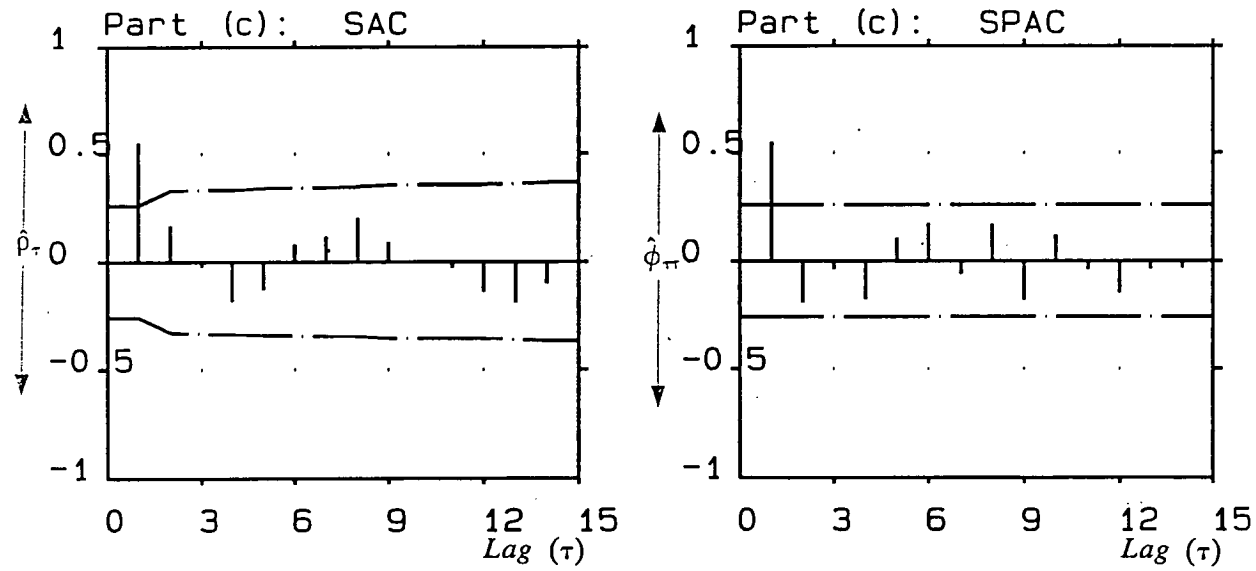
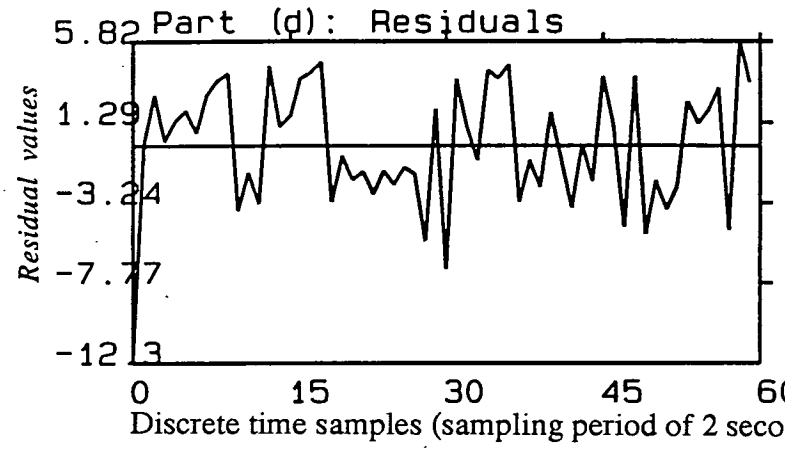


Figure 6.2: Part (a): Detrended averaged FHR (indirect phonocardiographic);
Part (b): Synthetic FHR data, generated from the parameters extracted from the boxed segment shown in part (a); the estimated parameter values were: $\hat{\phi}_1 = 0.8914$, $\hat{\phi}_2 = -0.1630$, $\hat{\sigma}_e = 3.3573$, $\hat{\sigma}_p = 5.2978$, $\hat{\mu}_{baseline} = 126.3513$; *Pseudo –periodicity* was not detected.
Part (c): SAC and SPAC of the first 60 synthetic data samples shown in part (b), indicating AR(1) correlation characteristics; The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.



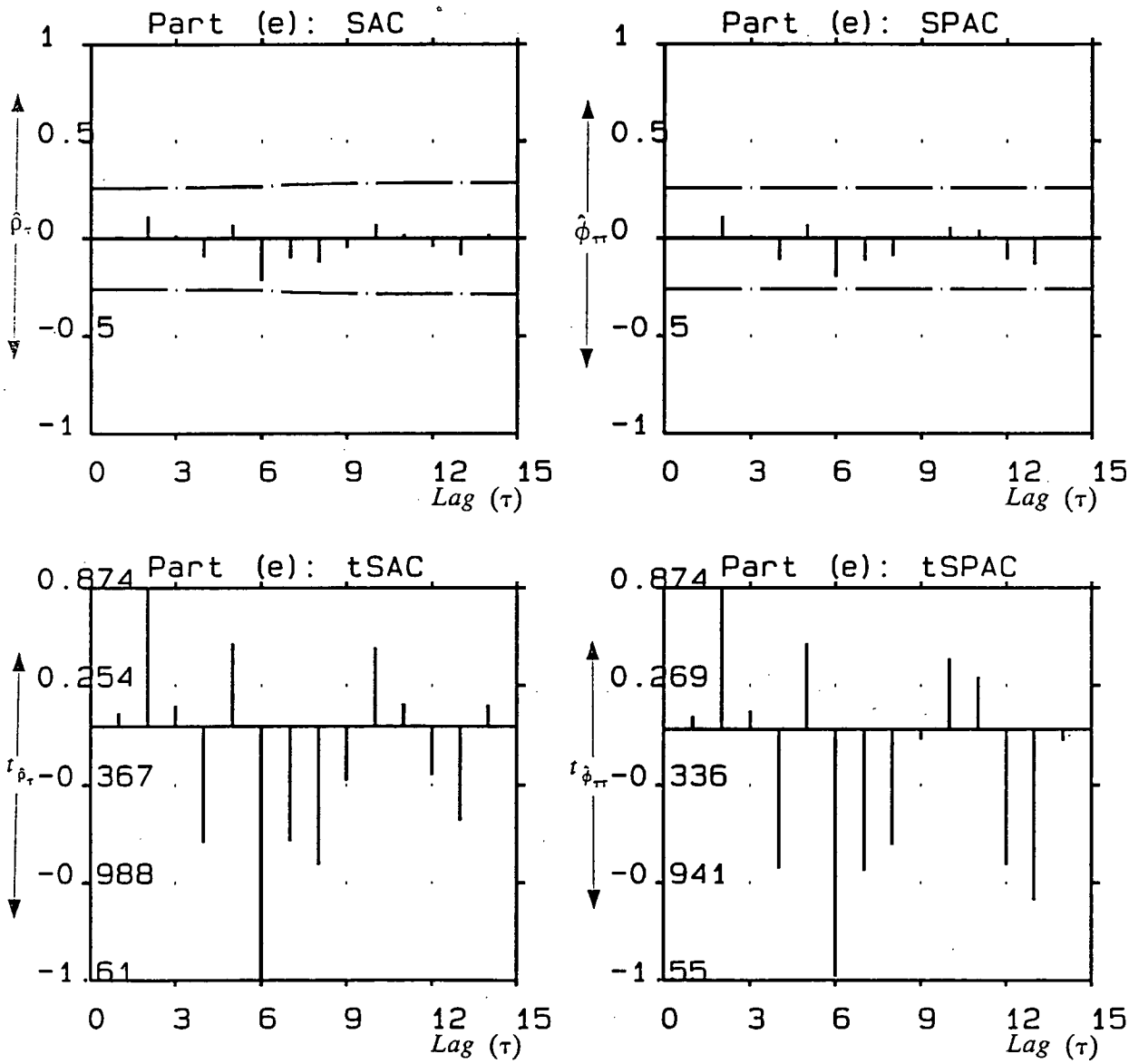
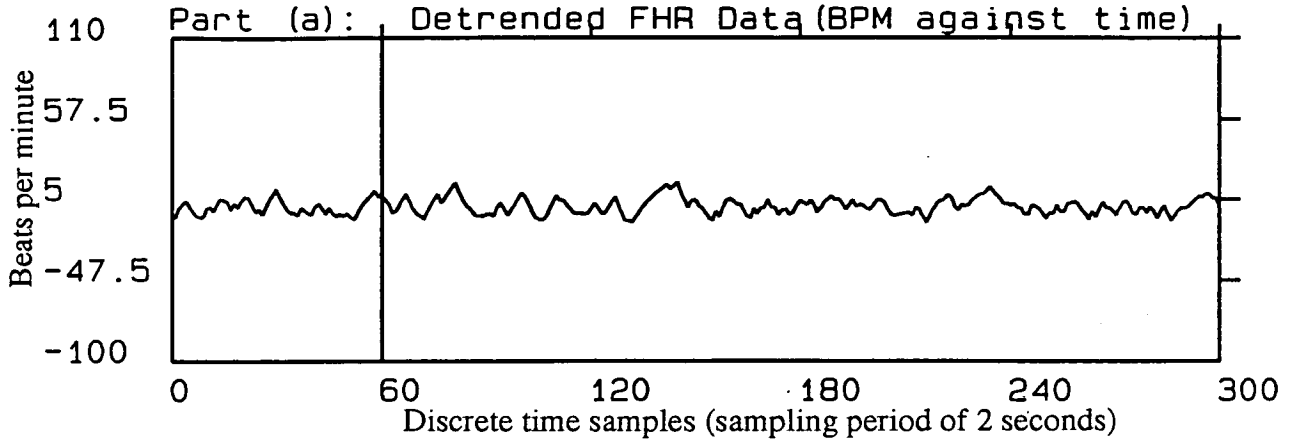
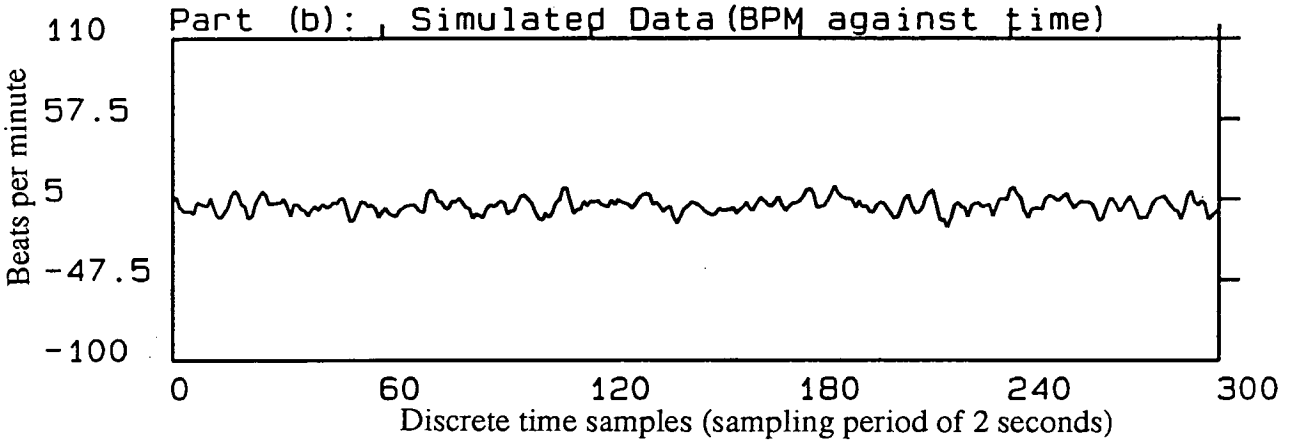


Figure 6.2: Part (d): Residuals (or Prediction errors or innovations) left after fitting an AR(2) model to the boxed segment shown in part (a).
Part (e): SAC, SPAC, t-statistics of SAC and SPAC of residuals shown above, indicating a white sequence; Q *–statistic* = 5.4246 $< \chi^2_{0.05}[13] = 22.362121$, further proof of the adequacy of the model fitted to the data.





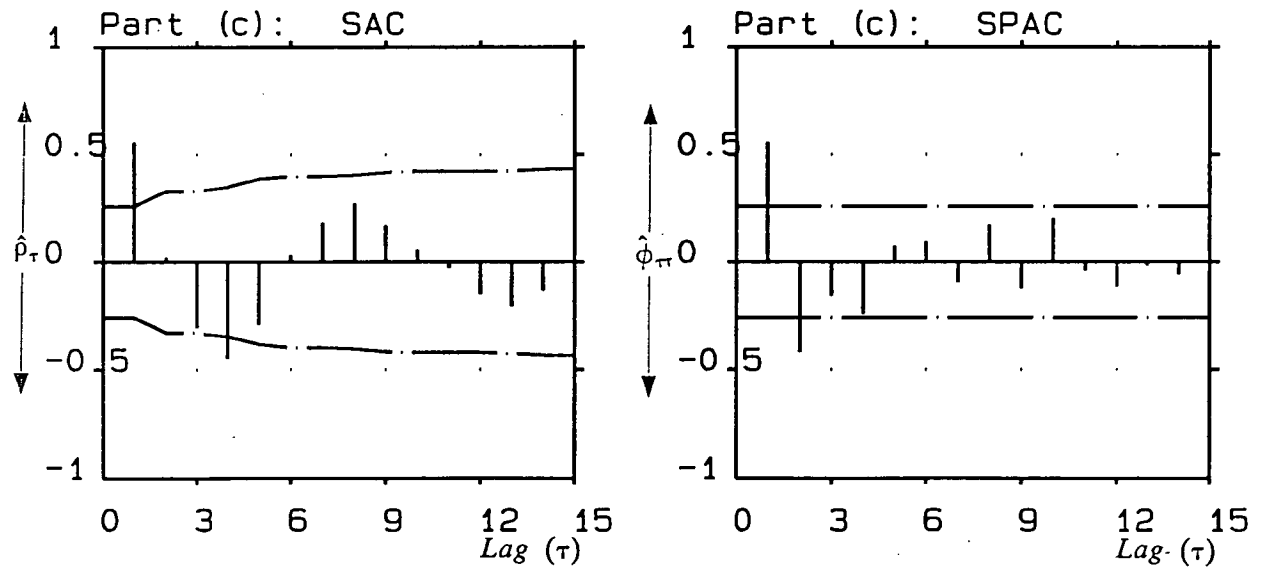
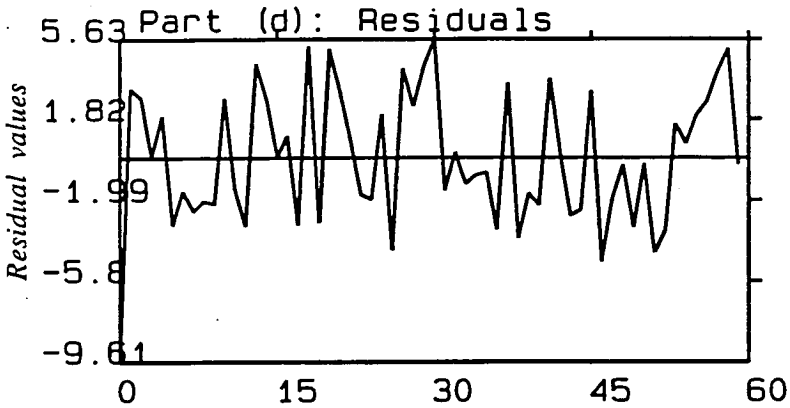


Figure 6.3: Part (a): Detrended averaged FHR (indirect phonocardiographic);
Part (b): Synthetic FHR data, generated from the parameters extracted from the boxed segment shown in part (a); the estimated parameter values were: $\hat{\phi}_1 = 0.9679$, $\hat{\phi}_2 = -0.3850$, $\hat{\sigma}_e = 3.1450$, $\hat{\sigma}_p = 4.7640$, $\hat{\mu}_{baseline} = 123.3177$; *Pseudo –periodicity* was detected ($t_o = 24.4188$).
Part (c): SAC and SPAC of the first 60 synthetic data samples shown in part (b), indicating AR(2) correlation characteristics; The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.



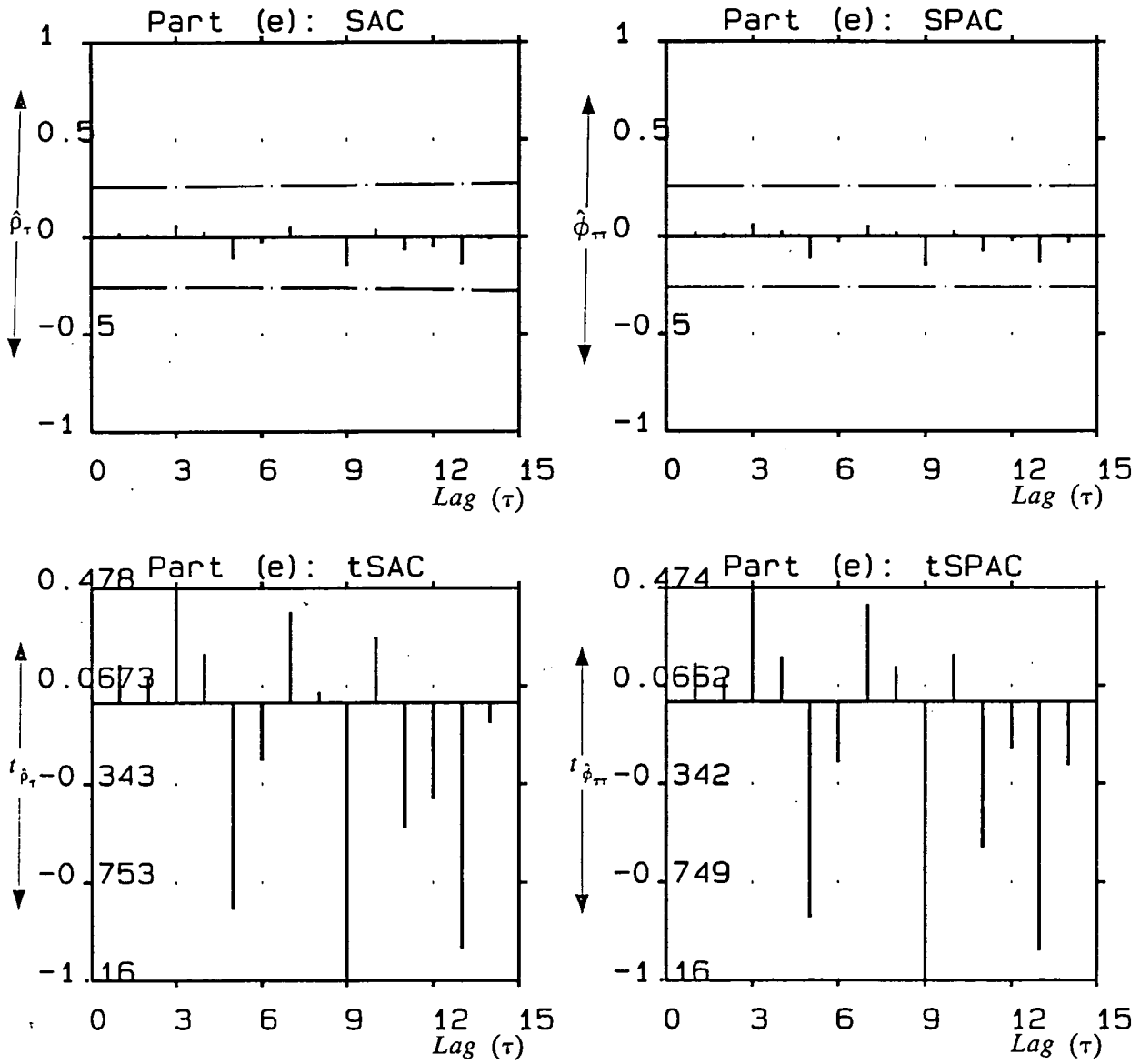


Figure 6.3: Part (d): Residuals (or Prediction errors or innovations) left after fitting an AR(2) model to the boxed segment shown in part (a).
Part (e): SAC, SPAC, t-statistics of SAC and SPAC of residuals shown above, indicating a white sequence; Q *–statistic* = 10.2924 $<\chi^2_{0.05}[13]=22.362121$, further proof of the adequacy of the model fitted to the data.

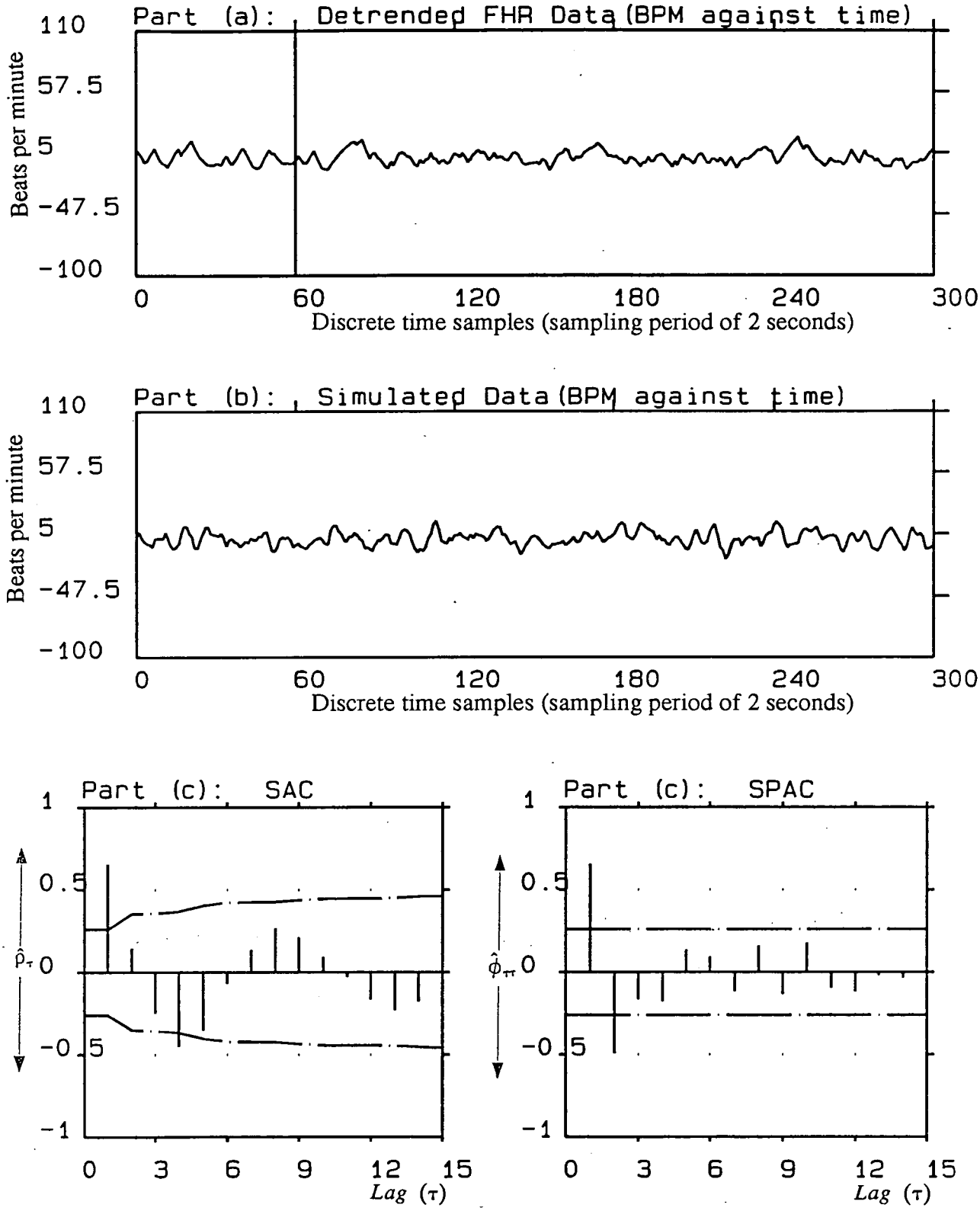
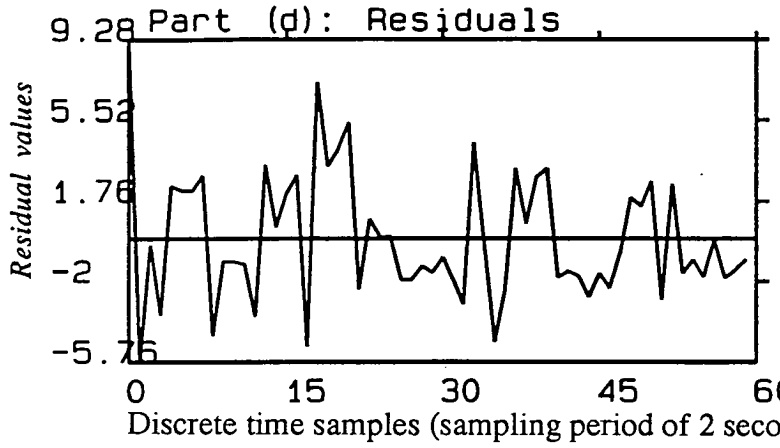


Figure 6.4: Part (a): Detrended averaged FHR (indirect phonocardiographic);
Part (b): Synthetic FHR data, generated from the parameters extracted from the boxed segment shown in part (a); the estimated parameter values were: $\hat{\phi}_1 = 1.1823$, $\hat{\phi}_2 = -0.4970$; $\hat{\sigma}_e = 2.9830$, $\hat{\sigma}_p = 5.6042$, $\hat{\mu}_{baseline} = 125.7053$; *Pseudo –periodicity* was detected ($t_o = 26.5820$).
Part (c): SAC and SPAC of the first 60 synthetic data samples shown in part (b), indicating AR(2) correlation characteristics; The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.



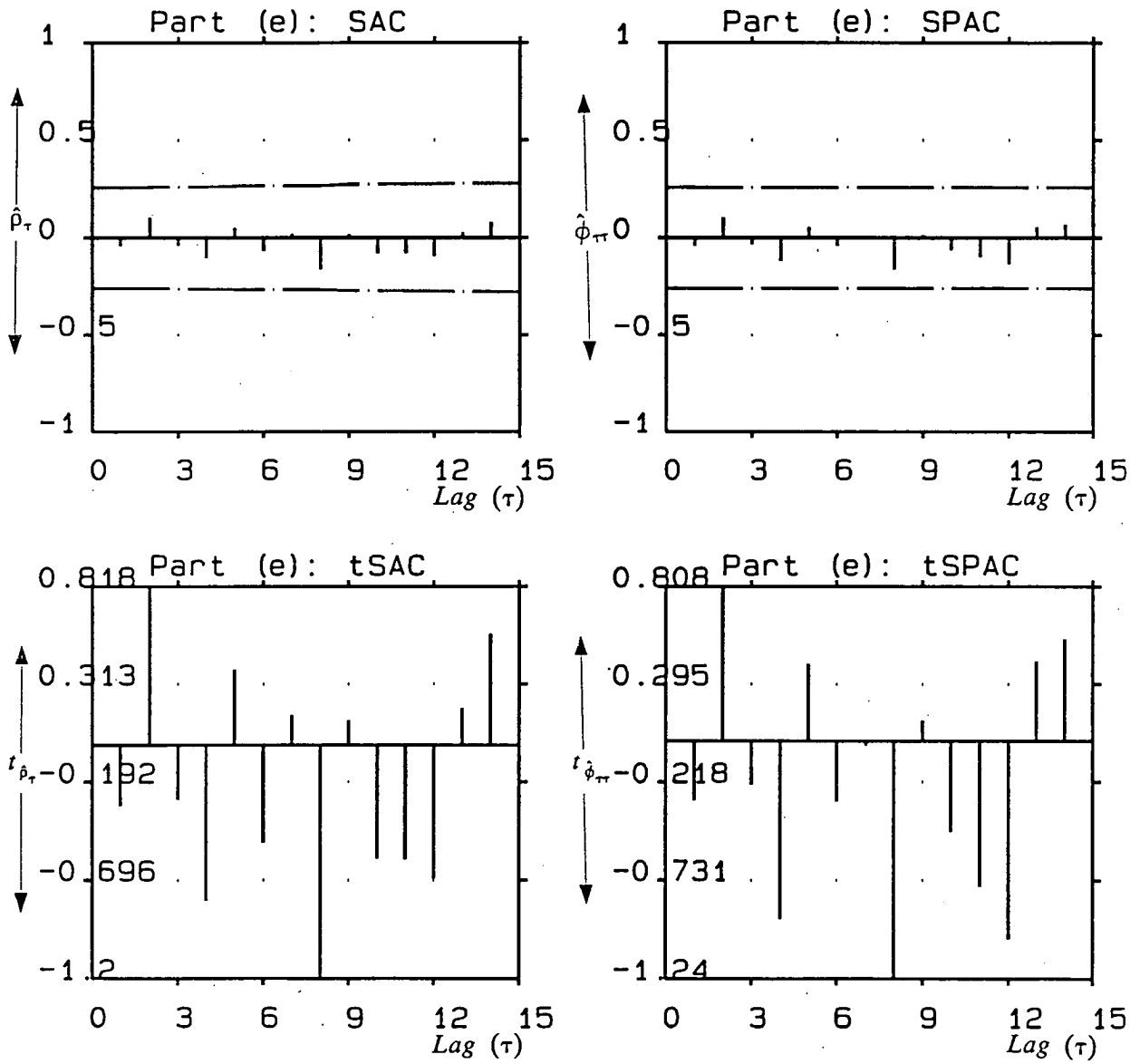
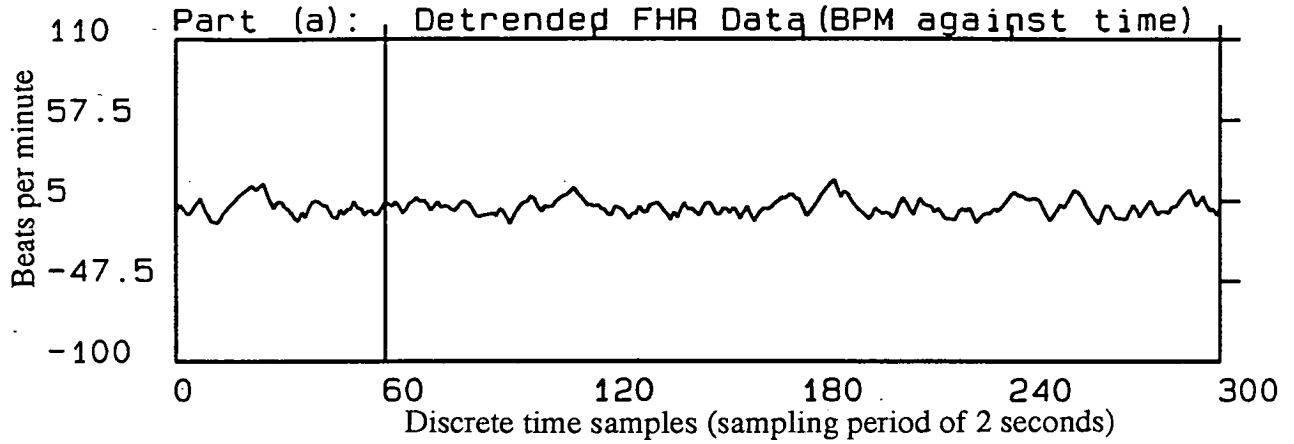
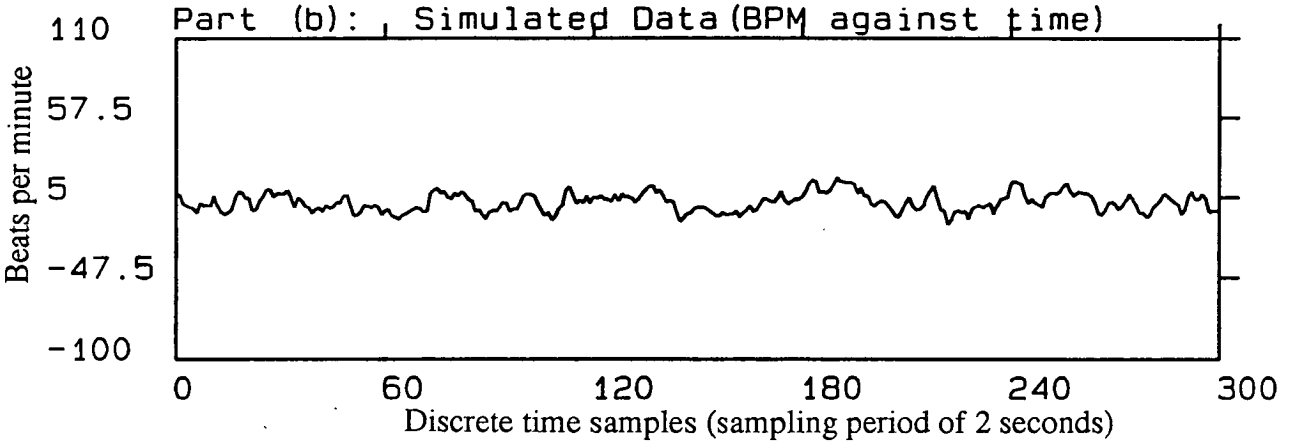


Figure 6.4: Part (d): Residuals (or Prediction errors or innovations) left after fitting an AR(2) model to the boxed segment shown in part (a).
Part (e): SAC, SPAC, t-statistics of SAC and SPAC of residuals shown above, indicating a white sequence; Q *−statistic* $= 6.4981 < \chi^2_{0.05}[13] = 22.3621$, further proof of the adequacy of the model fitted to the data.





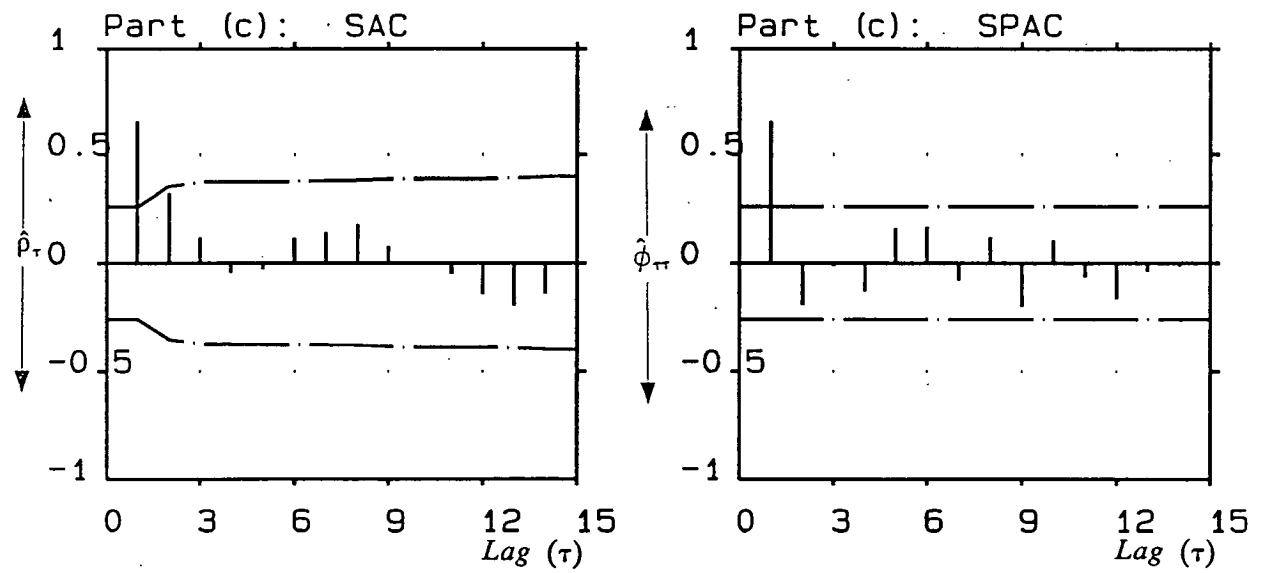


Figure 6.5: Part (a): Detrended averaged FHR (indirect phonocardiographic);
Part (b): Synthetic FHR data, generated from the parameters extracted from the boxed segment shown in part (a); the estimated parameter values were: $\hat{\phi}_1 = 1.0453$, $\hat{\phi}_2 = -0.2249$, $\hat{\sigma}_e = 3.1510$, $\hat{\sigma}_p = 6.2023$, $\hat{\mu}_{baseline} = 131.2902$; *Pseudo-periodicity* was not detected.
Part (c): SAC and SPAC of the first 60 synthetic data samples shown in part (b), indicating AR(1) correlation characteristics; The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.

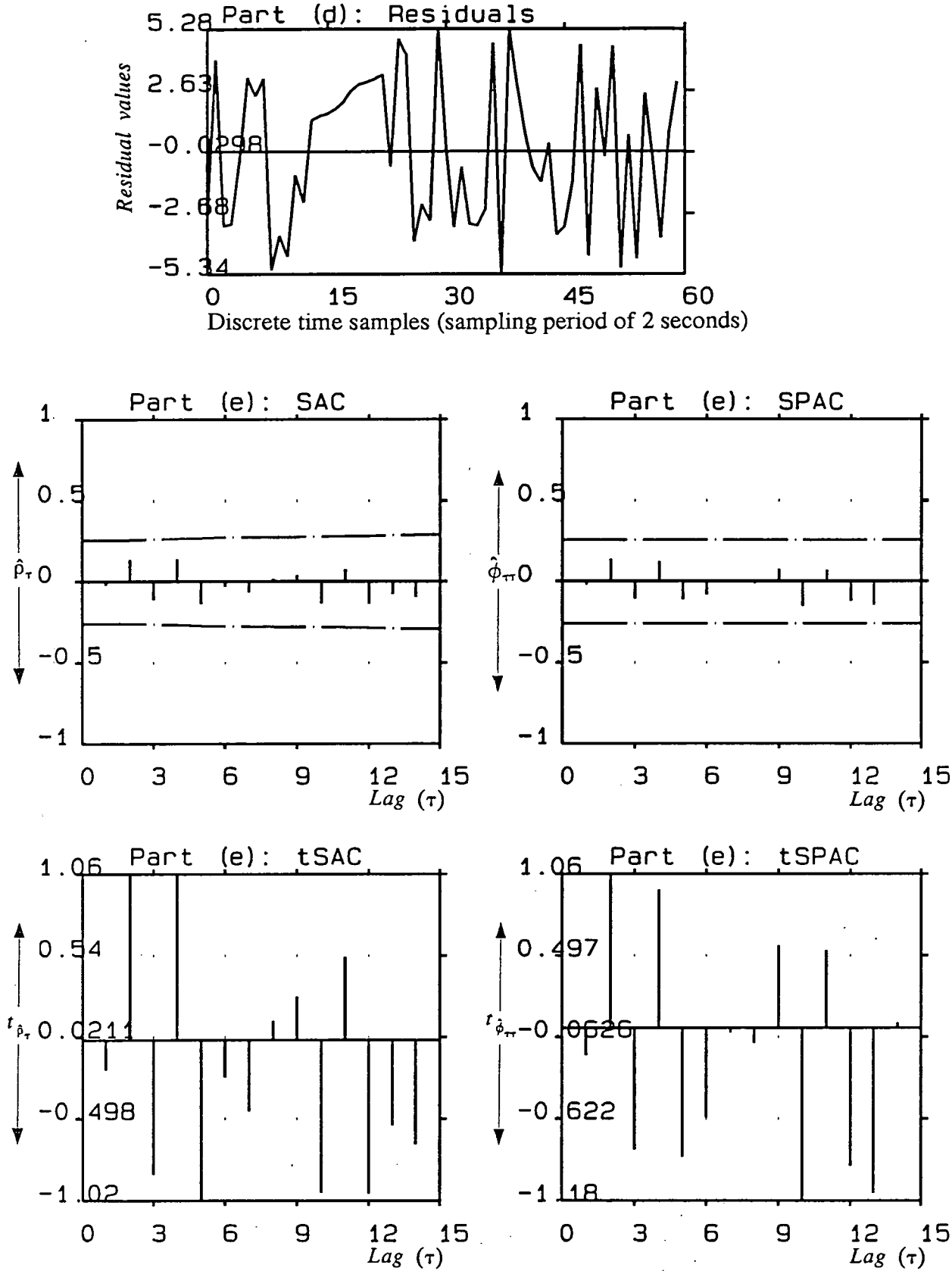


Figure 6.5: Part (d): Residuals (or Prediction errors or innovations) left after fitting an AR(2) model to the boxed segment shown in part (a).
Part (e): SAC, SPAC, t-statistics of SAC and SPAC of residuals shown above, indicating a white sequence; Q *–statistic* $= 9.5311 < \chi^2_{0.05}[13] = 22.3621$, further proof of the adequacy of the model fitted to the data.

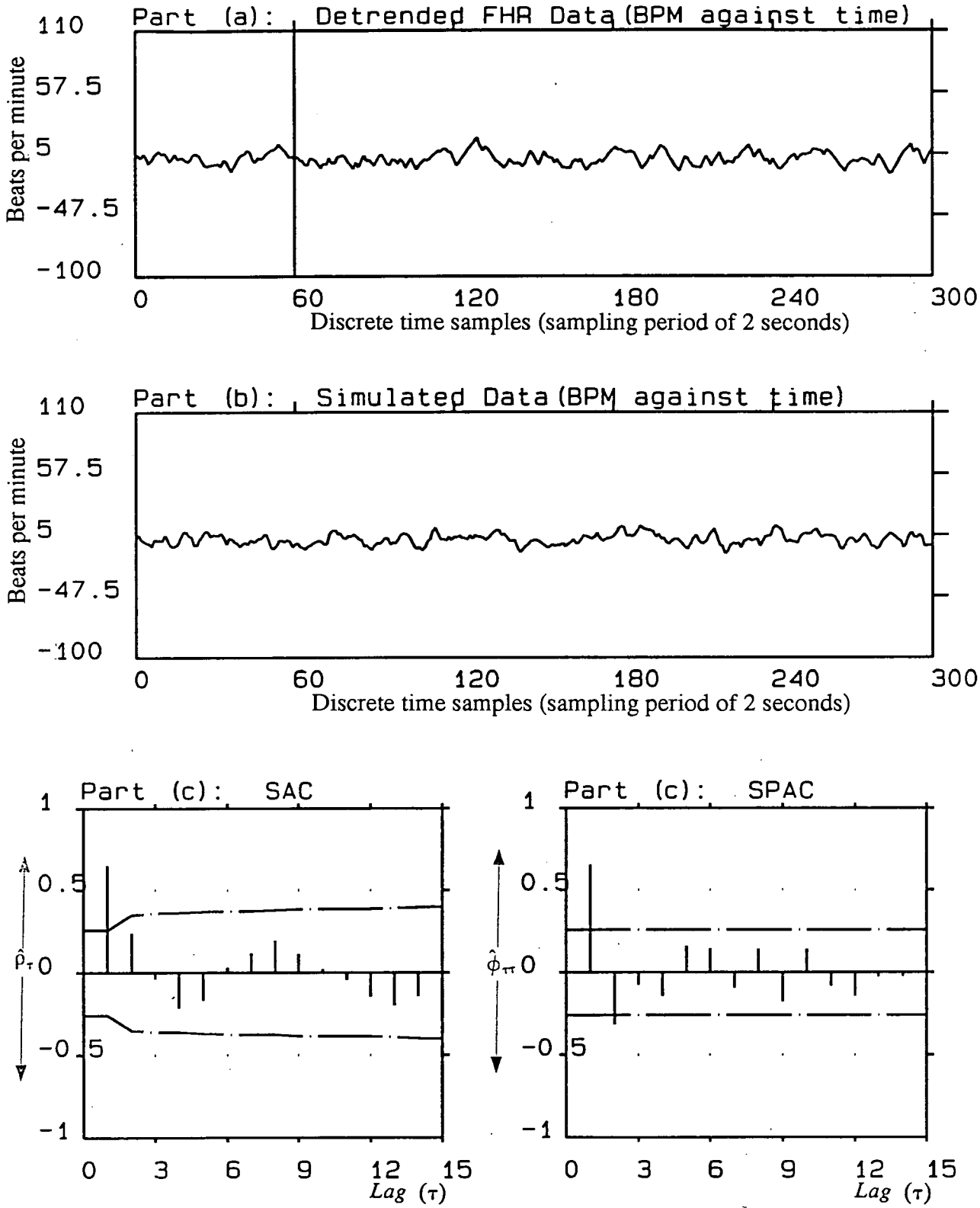
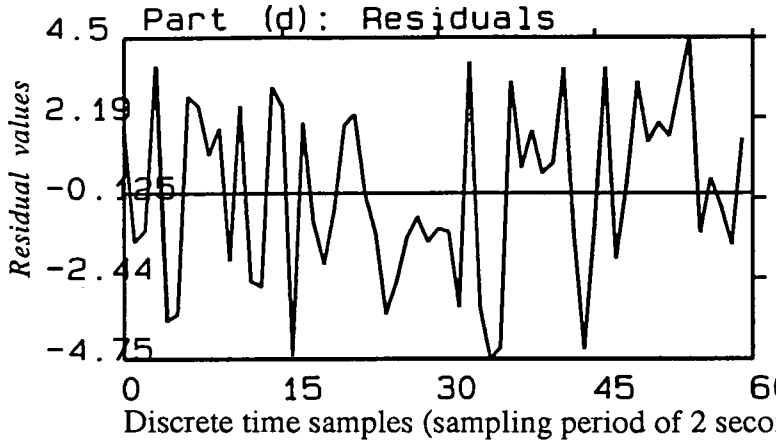


Figure 6.6: Part (a): Detrended averaged FHR (indirect phonocardiographic);
Part (b): Synthetic FHR data, generated from the parameters extracted from the boxed segment shown in part (a); the estimated parameter values were: $\hat{\phi}_1 = 1.1007$, $\hat{\phi}_2 = -0.3361$, $\hat{\sigma}_e = 2.4839$, $\hat{\sigma}_p = 4.6521$, $\hat{\mu}_{baseline} = 133.8820$; *Pseudo –periodicity* was not detected.
Part (c): SAC and SPAC of the first 60 synthetic data samples shown in part (b), indicating AR(2) correlation characteristics; The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.



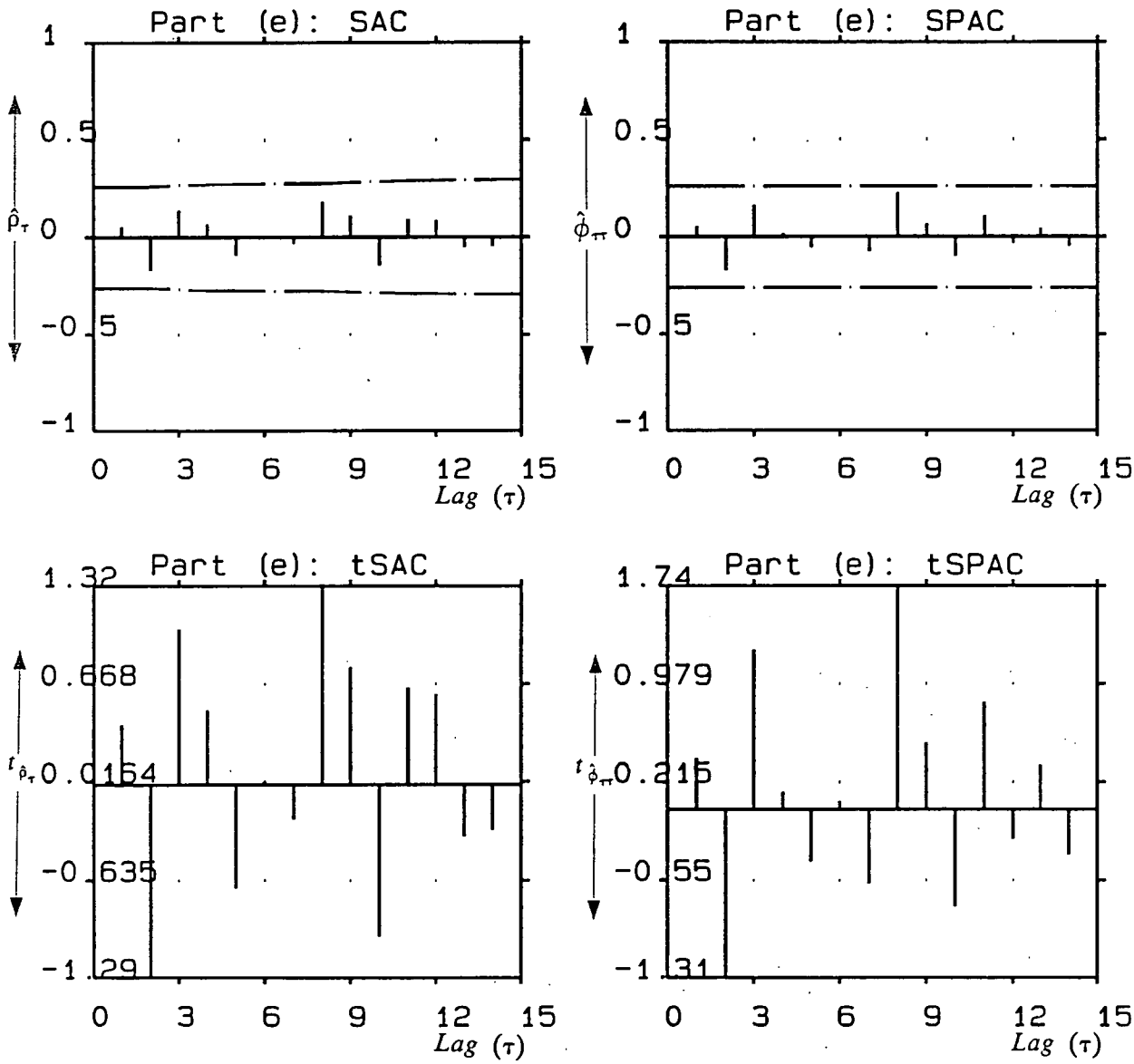
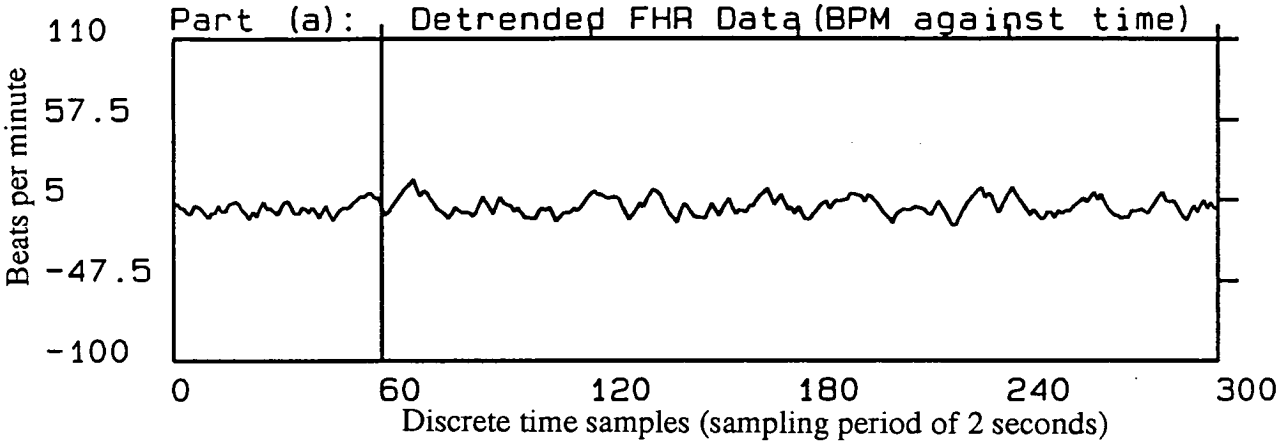
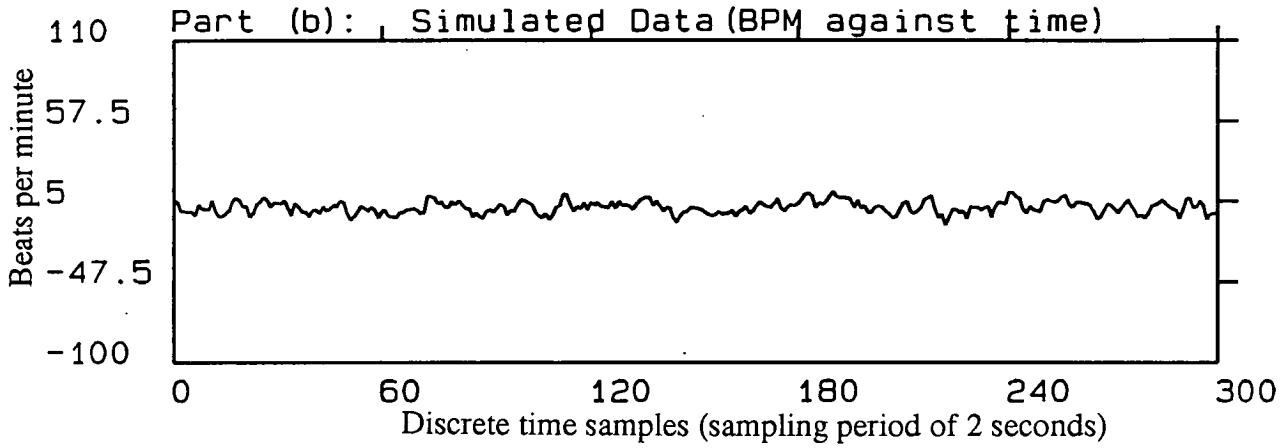


Figure 6.6: Part (d): Residuals (or Prediction errors or innovations) left after fitting an AR(2) model to the boxed segment shown in part (a).
Part (e): SAC, SPAC, t-statistics of SAC and SPAC of residuals shown above, indicating a white sequence; Q *-statistic* = 10.4951 $< \chi^2_{0.05}[13] = 22.3621$, further proof of the adequacy of the model fitted to the data.





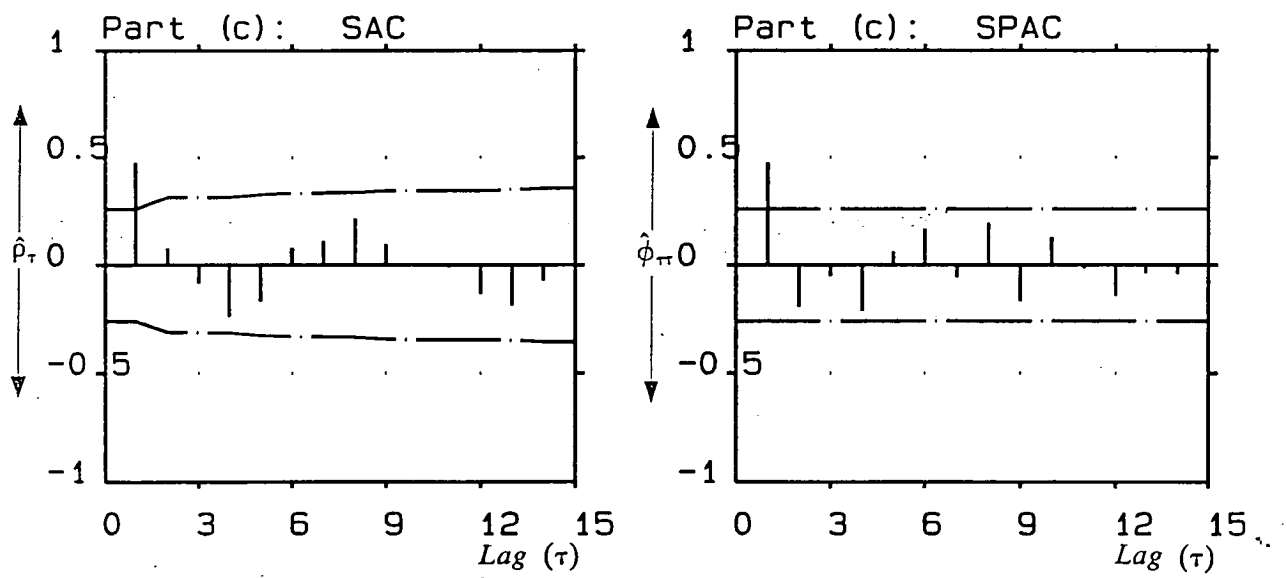
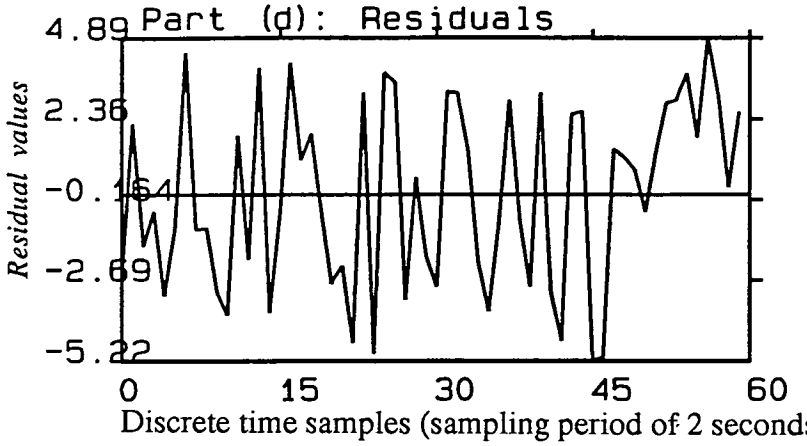


Figure 6.7: Part (a): Detrended averaged FHR (indirect phonocardiographic);
Part (b): Synthetic FHR data, generated from the parameters extracted from the boxed segment shown in part (a); the estimated parameter values were: $\hat{\phi}_1 = 0.7803$, $\hat{\phi}_2 = -0.1243$, $\hat{\sigma}_e = 2.8797$, $\hat{\sigma}_p = 4.0311$, $\hat{\mu}_{baseline} = 132.0989$; *Pseudo –periodicity* was not detected.
Part (c): SAC and SPAC of the first 60 synthetic data samples shown in part (b), indicating AR(1) correlation characteristics; The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.



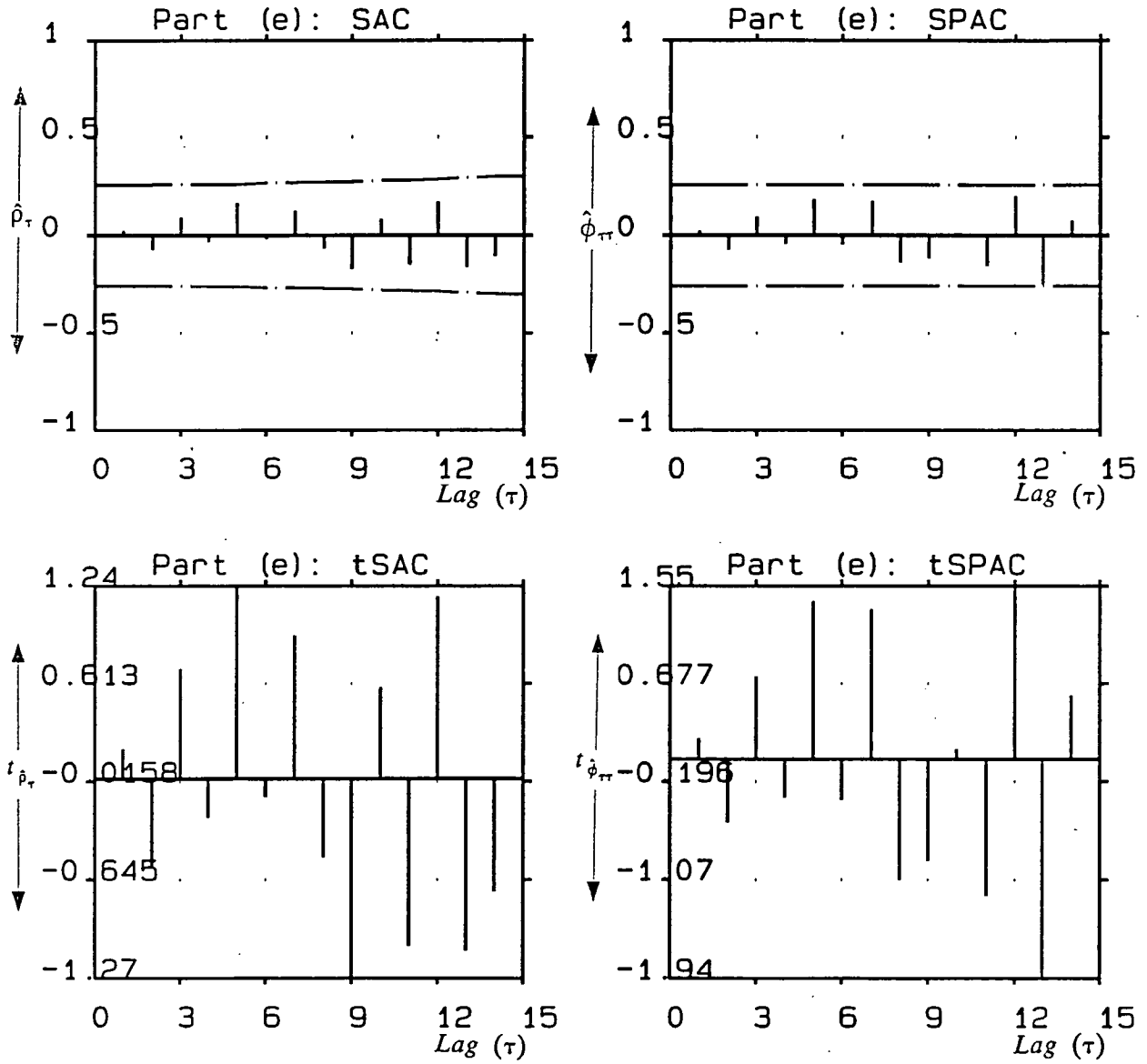
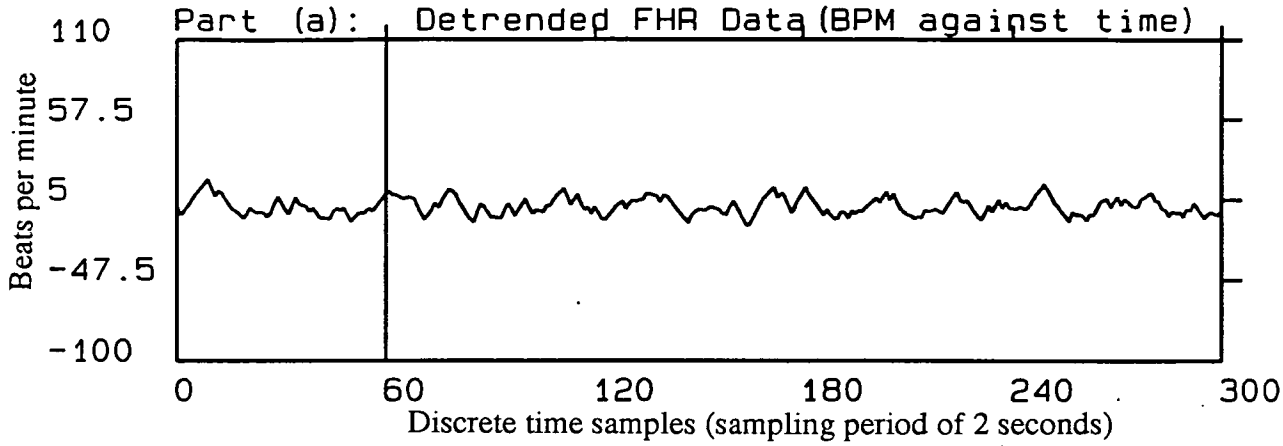
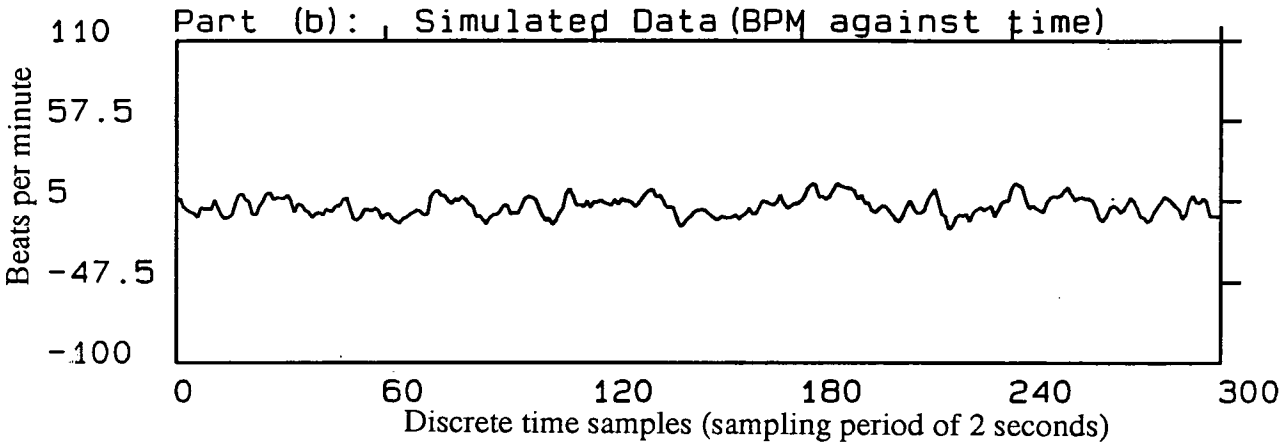


Figure 6.7: Part (d): Residuals (or Prediction errors or innovations) left after fitting an AR(2) model to the boxed segment shown in part (a).
Part (e): SAC, SPAC, t-statistics of SAC and SPAC of residuals shown above, indicating a white sequence; Q *–statistic* = 14.8523 $<\chi^2_{0.05}[13]=22.3621$, further proof of the adequacy of the model fitted to the data.





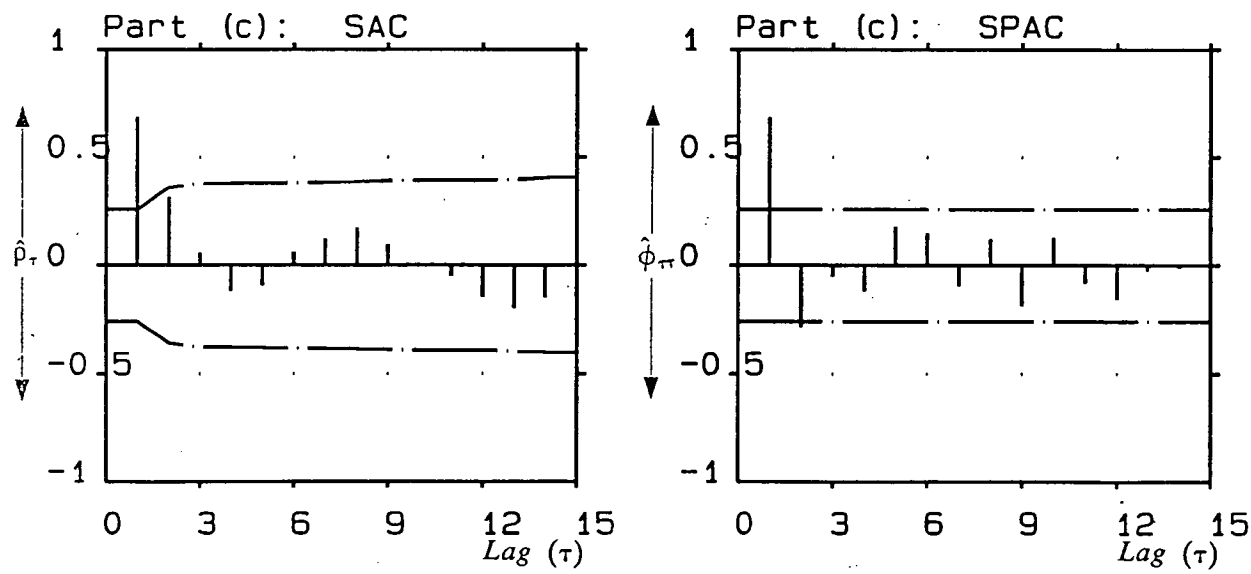
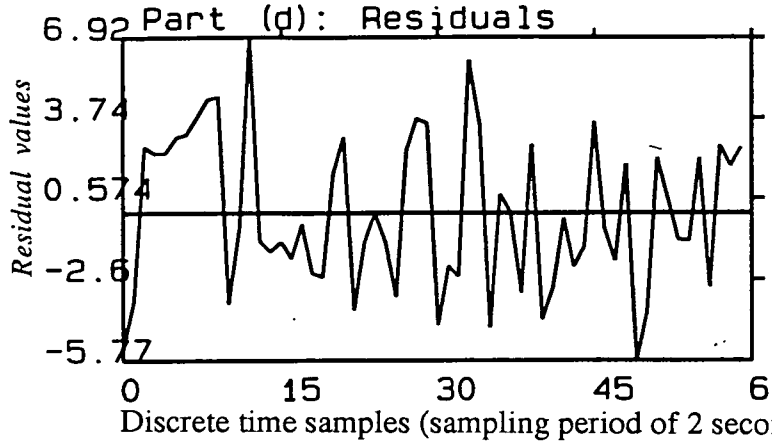


Figure 6.8: Part (a): Detrended averaged FHR (indirect phonocardiographic);
Part (b): Synthetic FHR data, generated from the parameters extracted from the boxed segment shown in part (a); the estimated parameter values were: $\hat{\phi}_1 = 1.1364$, $\hat{\phi}_2 = -0.3317$, $\hat{\sigma}_e = 2.9936$, $\hat{\sigma}_p = 6.0868$, $\hat{\mu}_{baseline} = 134.6735$; *Pseudo –periodicity* was not detected.
Part (c): SAC and SPAC of the first 60 synthetic data samples shown in part (b), indicating AR(2) correlation characteristics; The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.



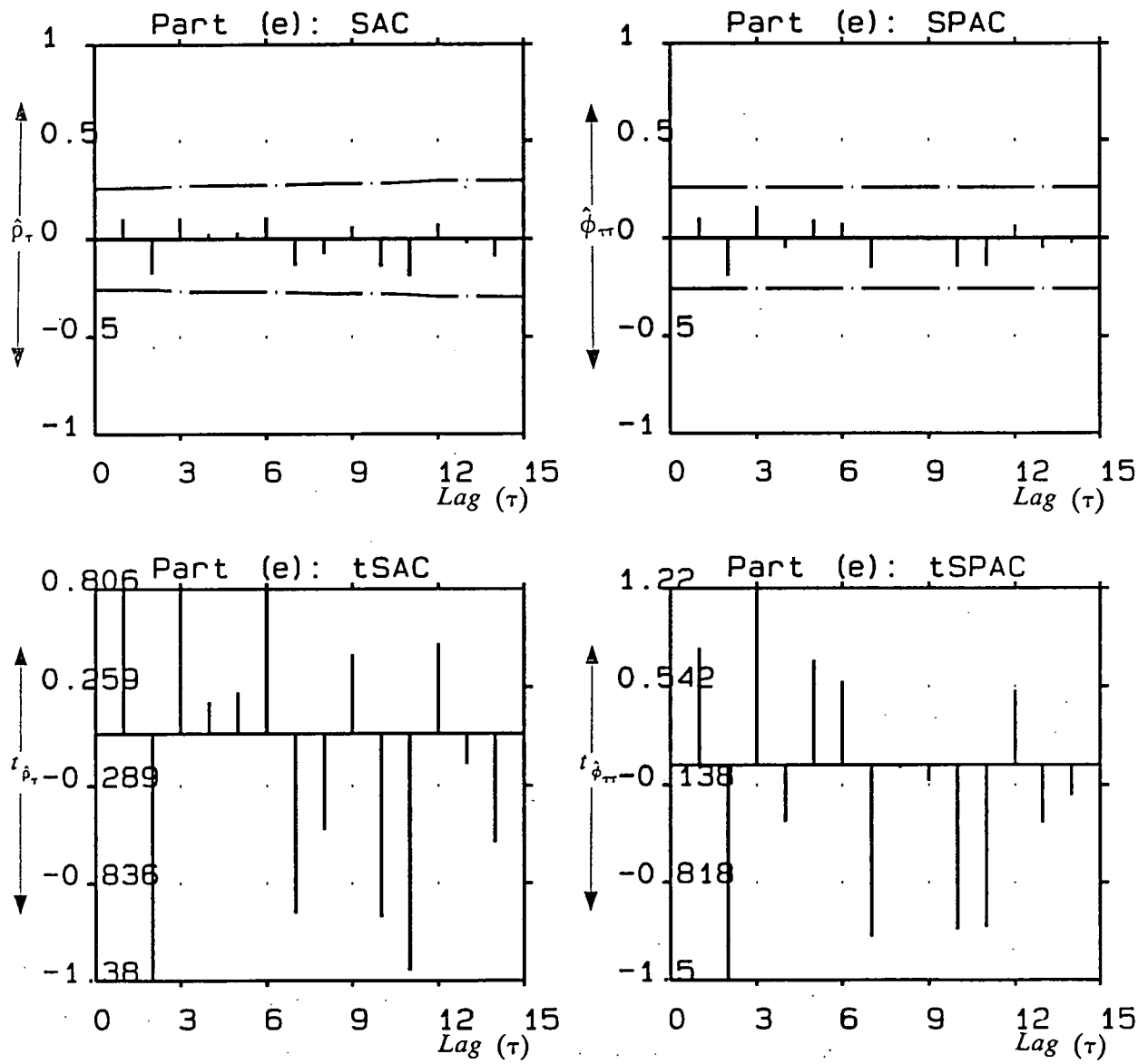
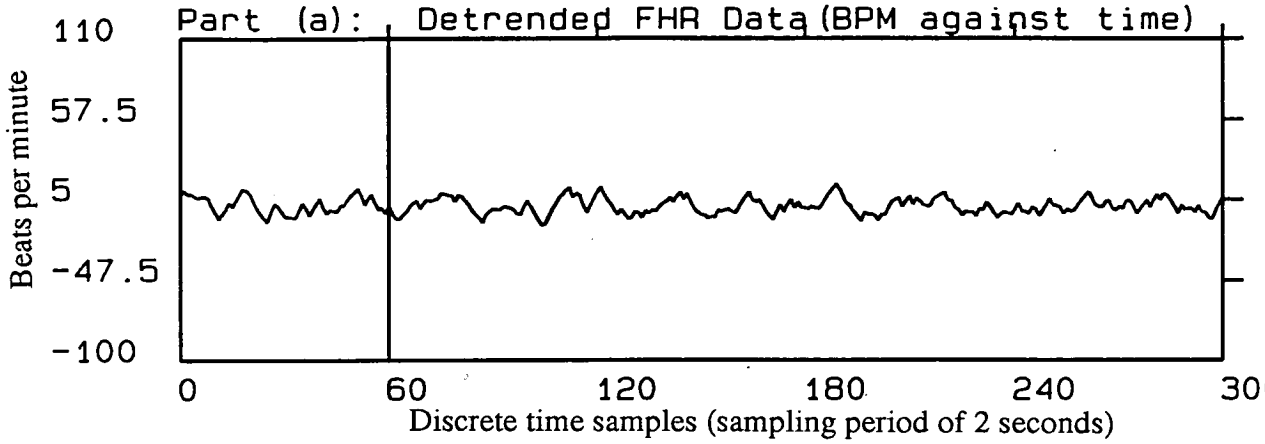
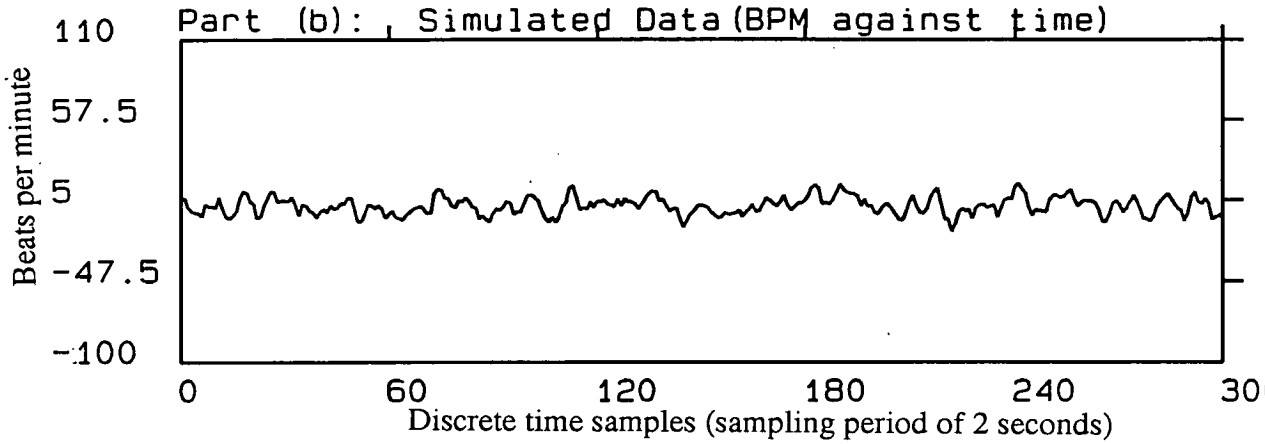


Figure 6.8: Part (d): Residuals (or Prediction errors or innovations) left after fitting an AR(2) model to the boxed segment shown in part (a).
Part (e): SAC, SPAC, t-statistics of SAC and SPAC of residuals shown above, indicating a white sequence; Q *–statistic* = 11.9666 $<\chi^2_{0.05}[13]=22.3621$, further proof of the adequacy of the model fitted to the data.





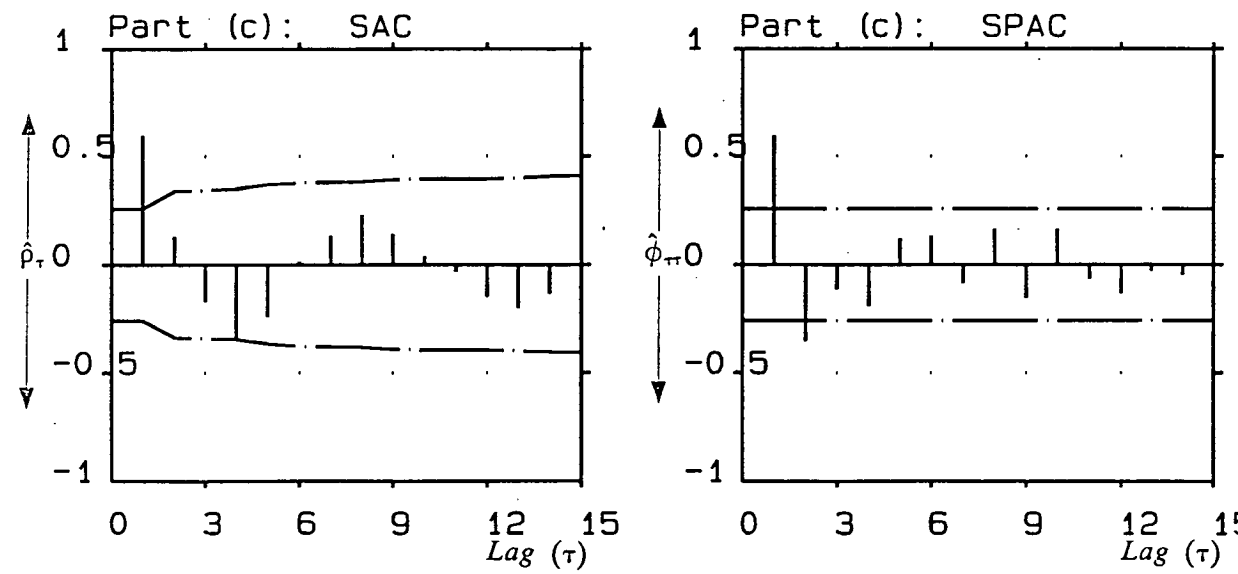


Figure 6.9: Part (a): Detrended averaged FHR (indirect phonocardiographic);
Part (b): Synthetic FHR data, generated from the parameters extracted from the boxed segment shown in part (a); the estimated parameter values were: $\hat{\phi}_1 = 1.0293$, $\hat{\phi}_2 = -0.3411$, $\hat{\sigma}_e = 3.4330$, $\hat{\sigma}_p = 5.6965$, $\hat{\mu}_{baseline} = 133.2532$; *Pseudo –periodicity* was not detected.
Part (c): SAC and SPAC of the first 60 synthetic data samples shown in part (b), indicating AR(2) correlation characteristics; The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.

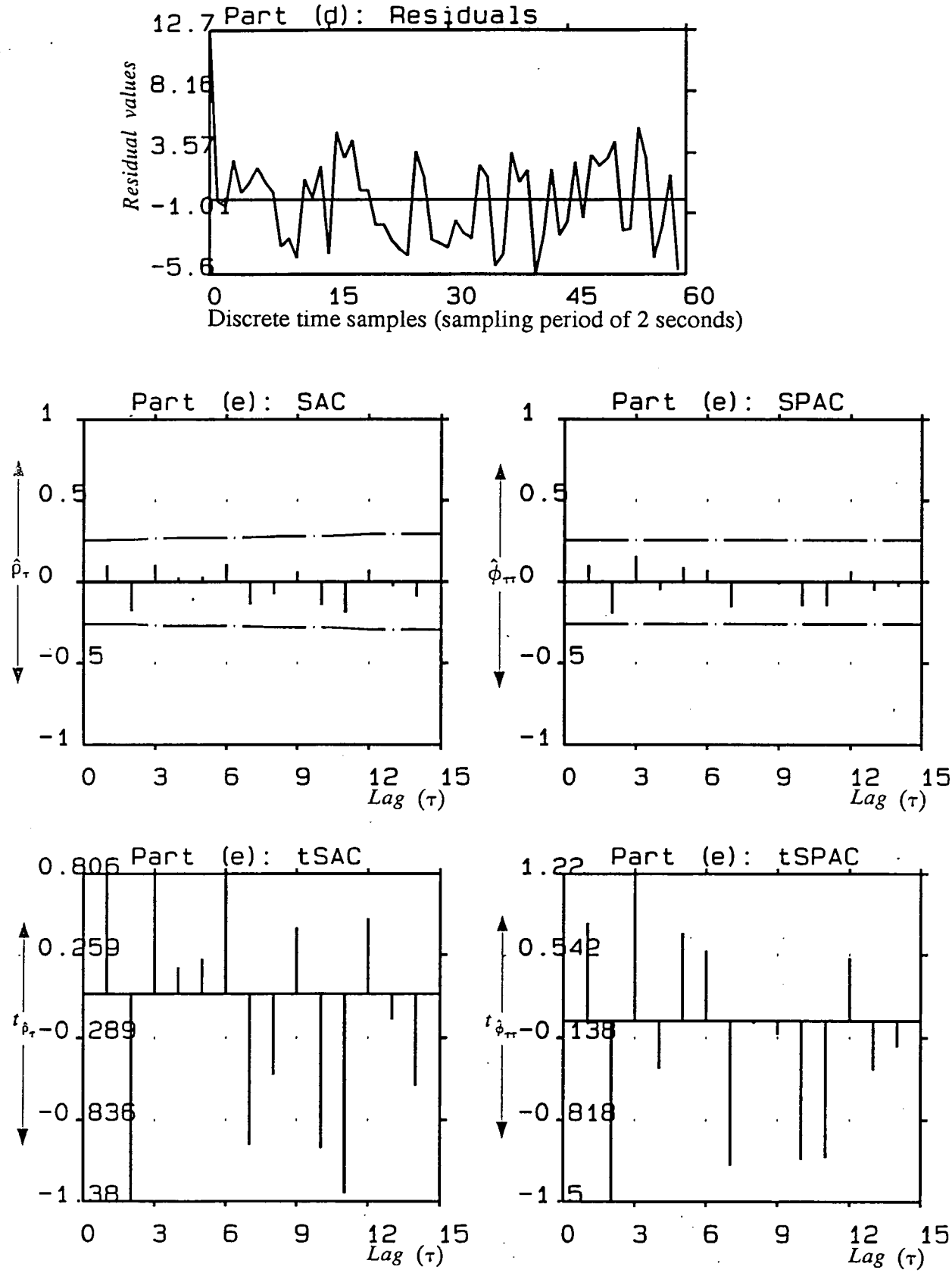
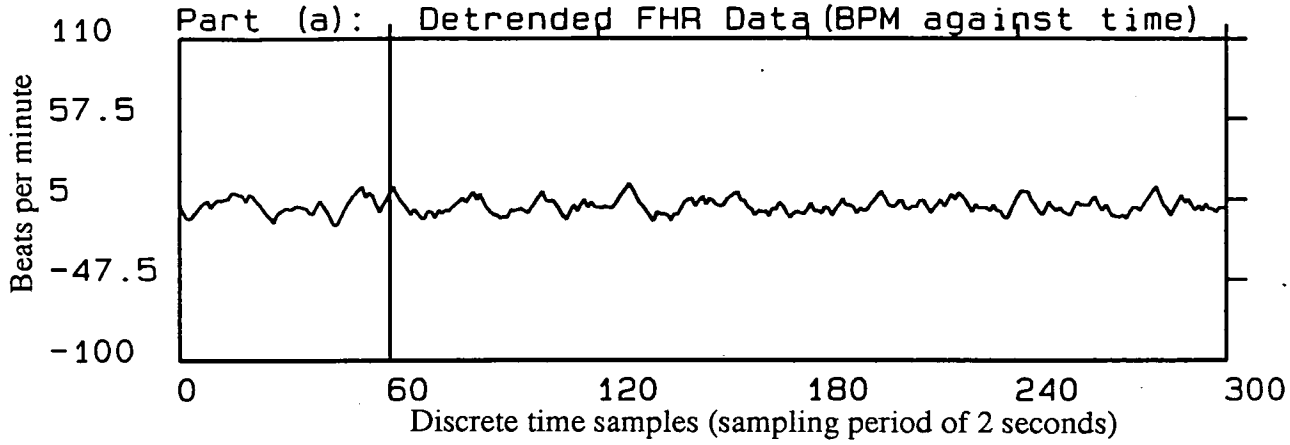
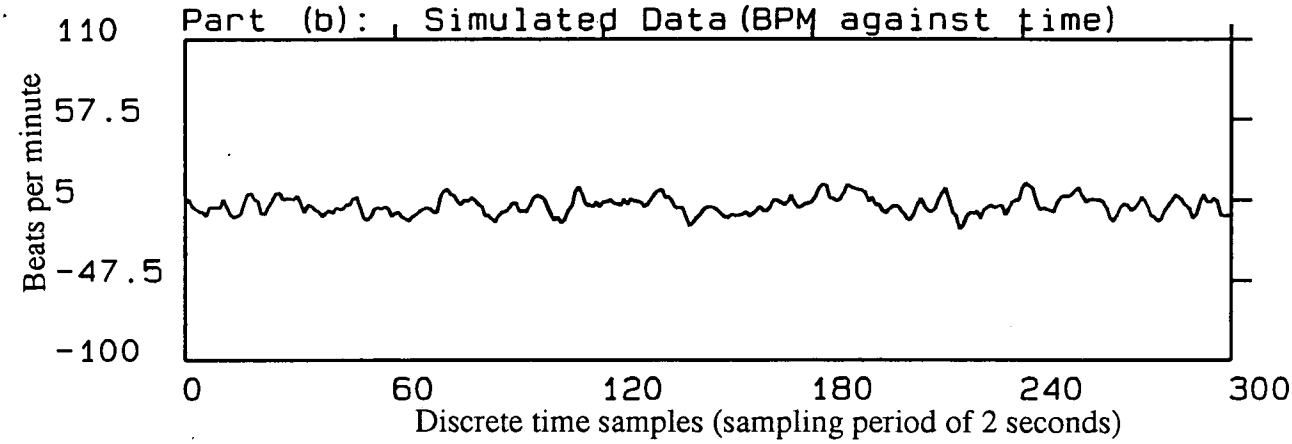


Figure 6.9: Part (d): Residuals (or Prediction errors or innovations) left after fitting an AR(2) model to the boxed segment shown in part (a).
Part (e): SAC, SPAC, t-statistics of SAC and SPAC of residuals shown above, indicating a white sequence; Q *-statistic* $= 7.7820 < \chi^2_{0.05}[13] = 22.3621$, further proof of the adequacy of the model fitted to the data.





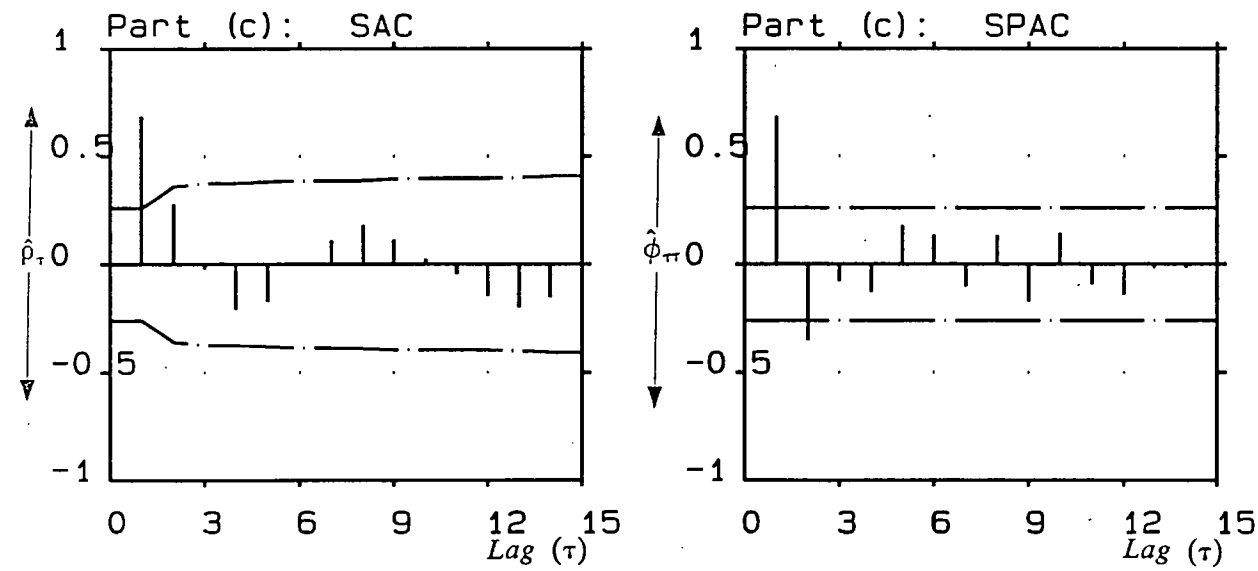
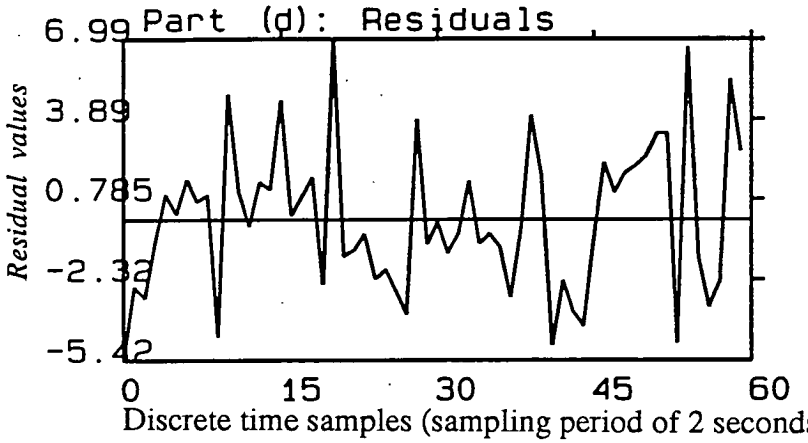


Figure 6.10: Part (a): Detrended averaged FHR (indirect phonocardiographic);
Part (b): Synthetic FHR data, generated from the parameters extracted from the boxed segment shown in part (a); the estimated parameter values were: $\hat{\phi}_1 = 1.1733$, $\hat{\phi}_2 = -0.3938$, $\hat{\sigma}_e = 2.8519$, $\hat{\sigma}_p = 5.7476$, $\hat{\mu}_{baseline} = 132.2469$; *Pseudo –periodicity* was not detected.
Part (c): SAC and SPAC of the first 60 synthetic data samples shown in part (b), indicating AR(2) correlation characteristics; The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.



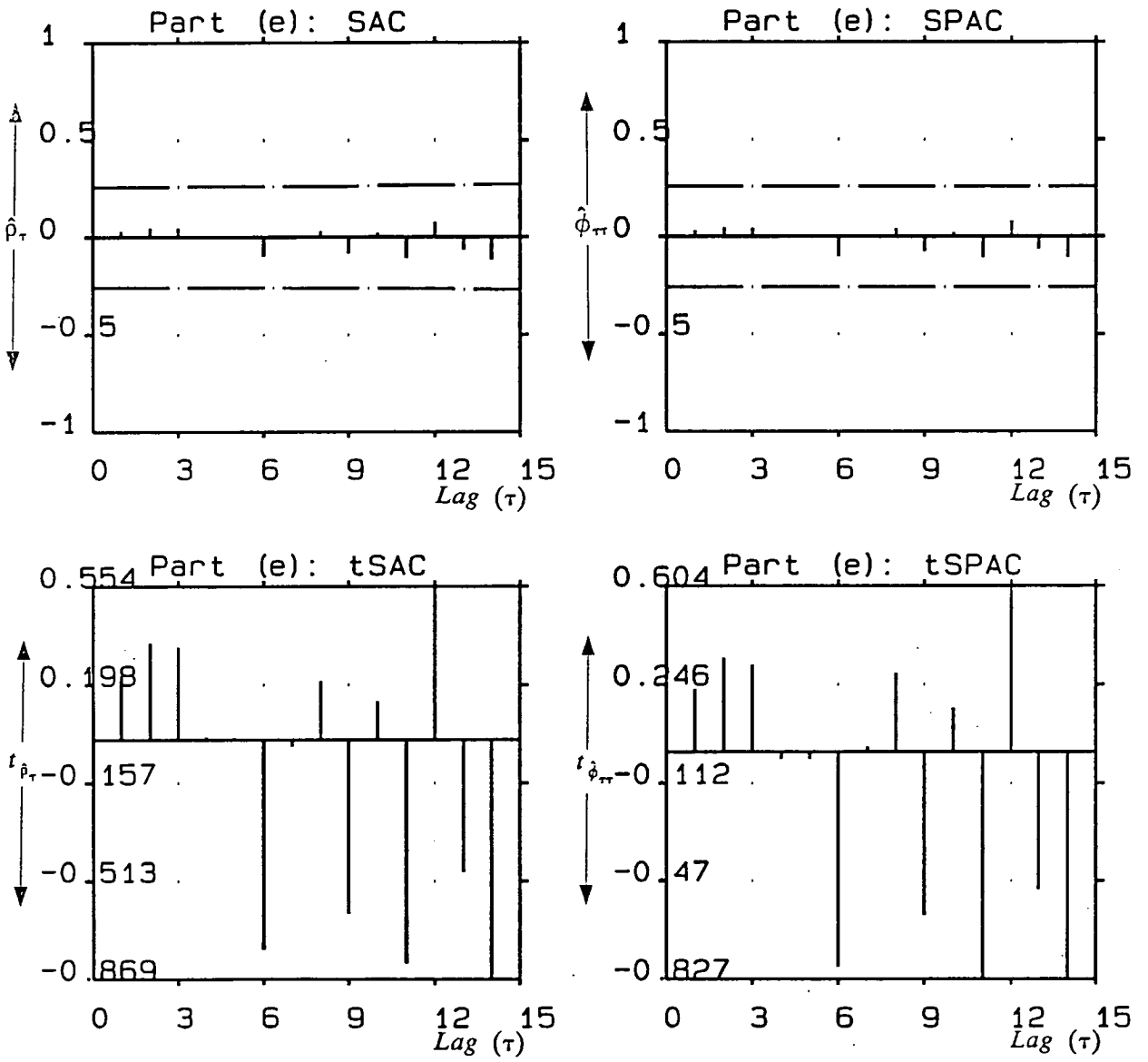


Figure 6.10: Part (d): Residuals (or Prediction errors or innovations) left after fitting an AR(2) model to the boxed segment shown in part (a).
Part (e): SAC, SPAC, t-statistics of SAC and SPAC of residuals shown above, indicating a white sequence; Q *–statistic* $= 4.5172 < \chi^2_{0.05}[13] = 22.3621$, further proof of the adequacy of the model fitted to the data.

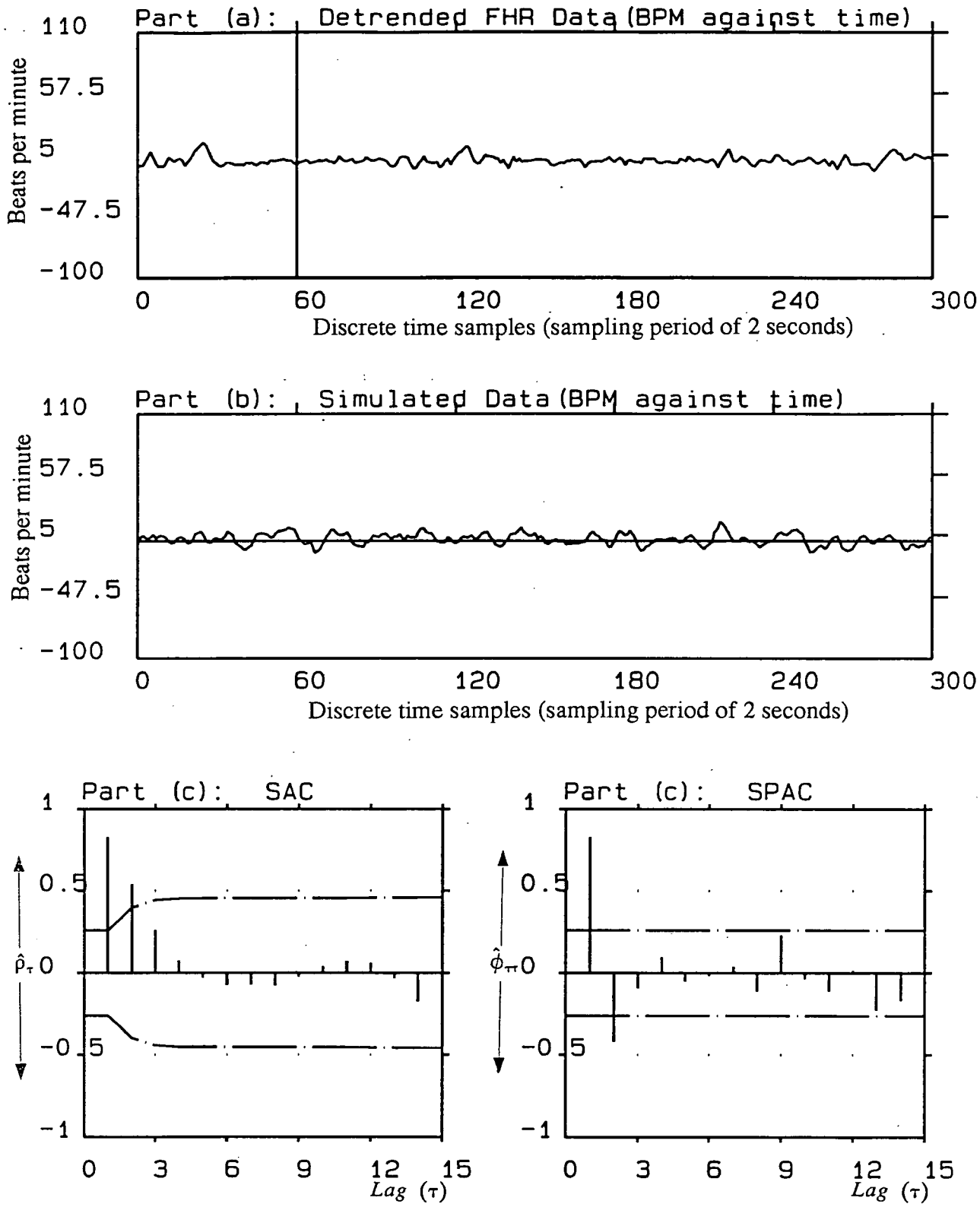
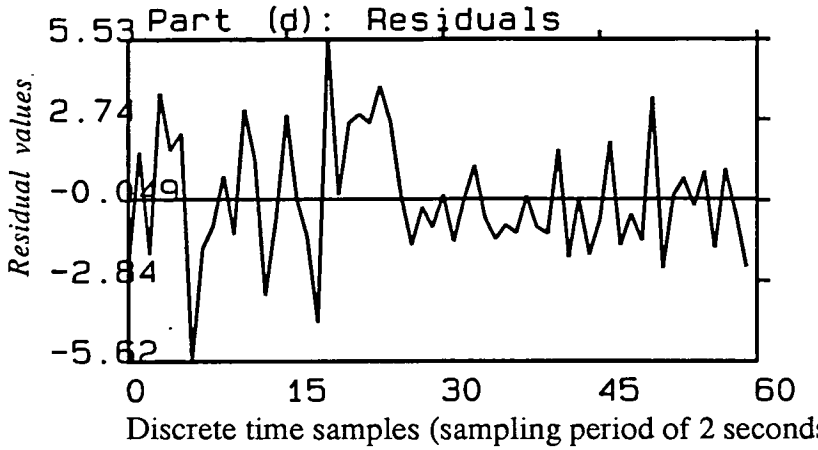


Figure 6.11: Part (a): Detrended averaged FHR (indirect phonocardiographic);
Part (b): Synthetic FHR data, generated from the parameters extracted from the boxed segment shown in part (a); the estimated parameter values were: $\hat{\phi}_1 = 1.3072$, $\hat{\phi}_2 = -0.5155$, $\hat{\sigma}_e = 2.0834$, $\hat{\sigma}_p = 4.8052$, $\hat{\mu}_{baseline} = 117.4582$; *Pseudo –periodicity* was detected ($t_o = 44.7029$).
Part (c): SAC and SPAC of the first 60 synthetic data samples shown in part (b), indicating AR(2) correlation characteristics; The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.



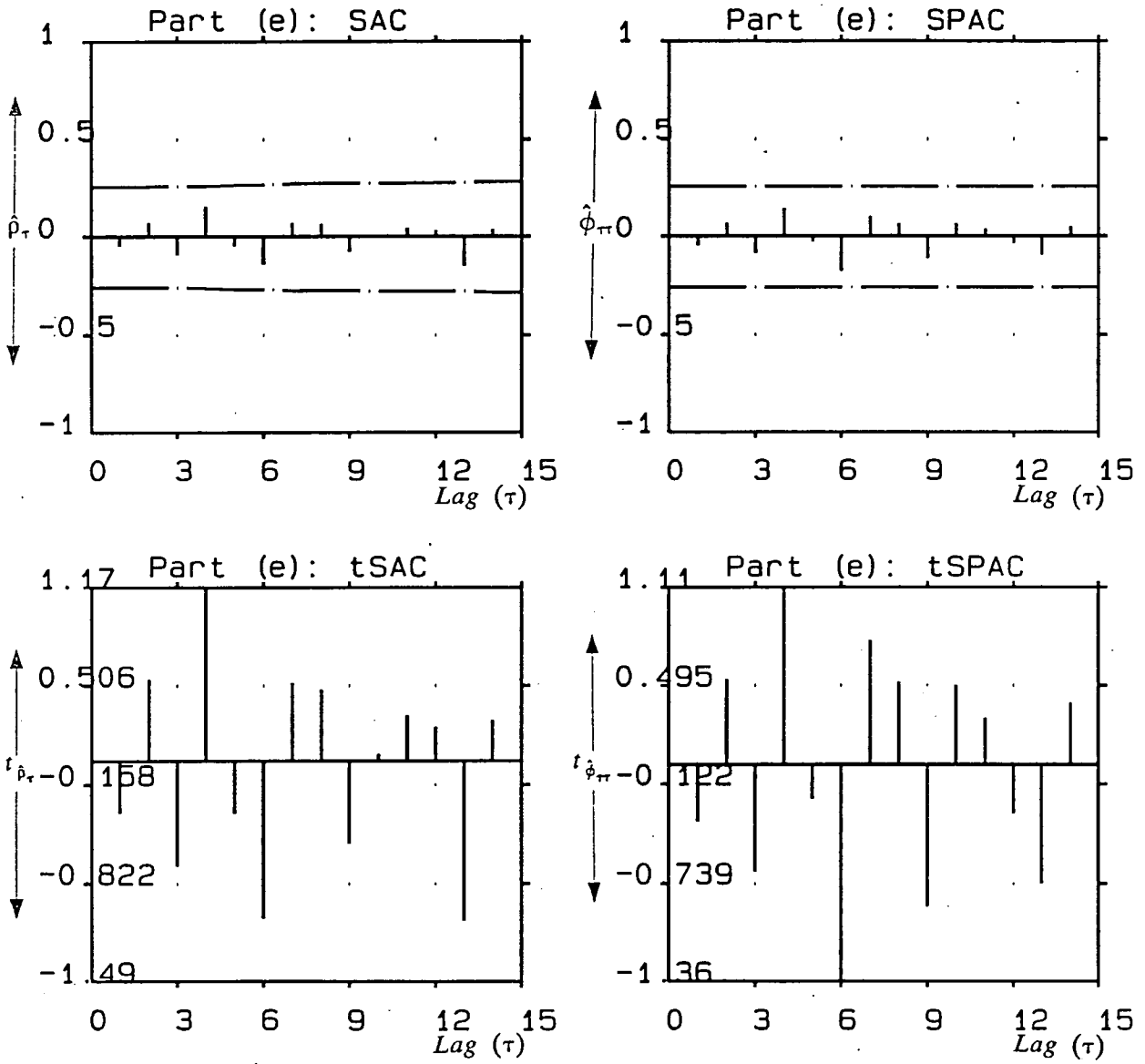


Figure 6.11: Part (d): Residuals (or Prediction errors or innovations) left after fitting an AR(2) model to the boxed segment shown in part (a).
Part (e): SAC, SPAC, t-statistics of SAC and SPAC of residuals shown above, indicating a white sequence; Q *–statistic* = 10.9099 $< \chi^2_{0.05}[13] = 22.3621$, further proof of the adequacy of the model fitted to the data.

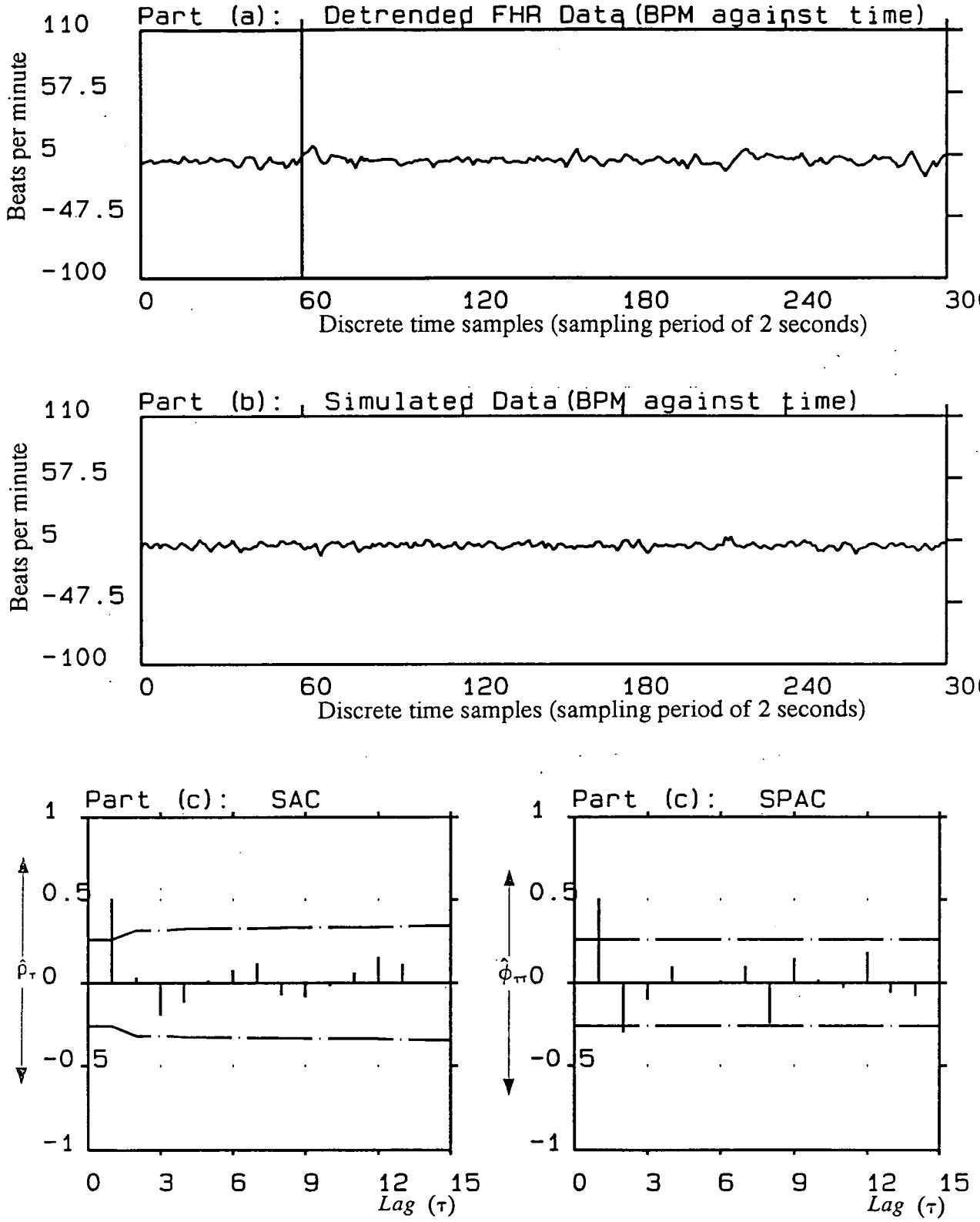
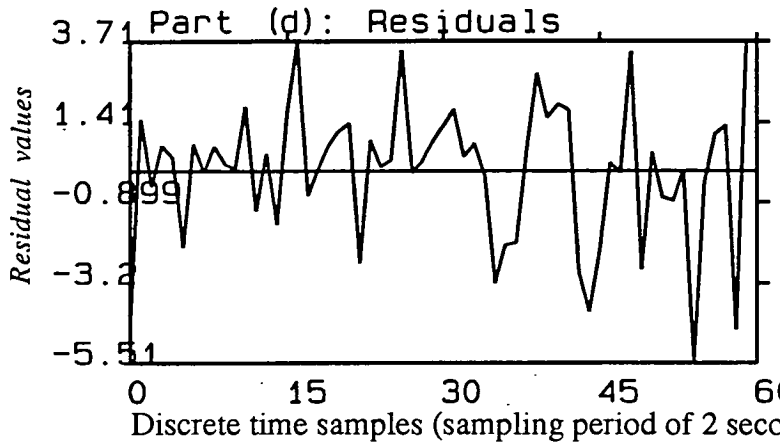


Figure 6.12: Part (a): Detrended averaged FHR (indirect phonocardiographic);
Part (b): Synthetic FHR data, generated from the parameters extracted from the boxed segment shown in part (a); the estimated parameter values were: $\hat{\phi}_1 = 0.6755$, $\hat{\phi}_2 = -0.2837$, $\hat{\sigma}_e = 1.9854$, $\hat{\sigma}_p = 2.4348$, $\hat{\mu}_{baseline} = 116.8660$; *Pseudo –periodicity* was detected ($t_o = 17.9270$).
Part (c): SAC and SPAC of the first 60 synthetic data samples shown in part (b), indicating AR(2) correlation characteristics; The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.



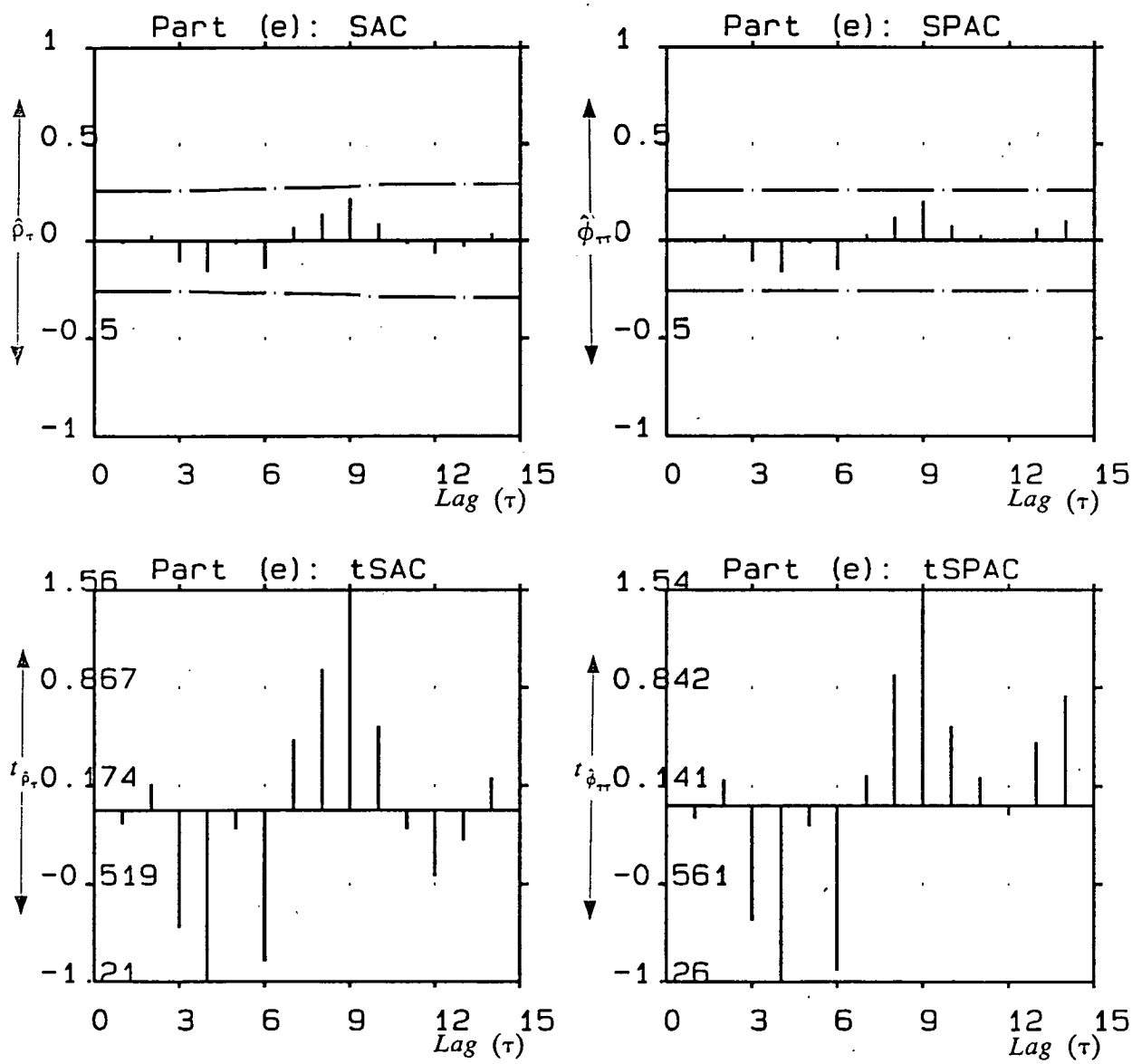


Figure 6.12: Part (d): Residuals (or Prediction errors or innovations) left after fitting an AR(2) model to the boxed segment shown in part (a).
Part (e): SAC, SPAC, t-statistics of SAC and SPAC of residuals shown above, indicating a white sequence; Q *–statistic* = 10.0651 $<\chi^2_{0.05}[13]=22.3621$, further proof of the adequacy of the model fitted to the data.

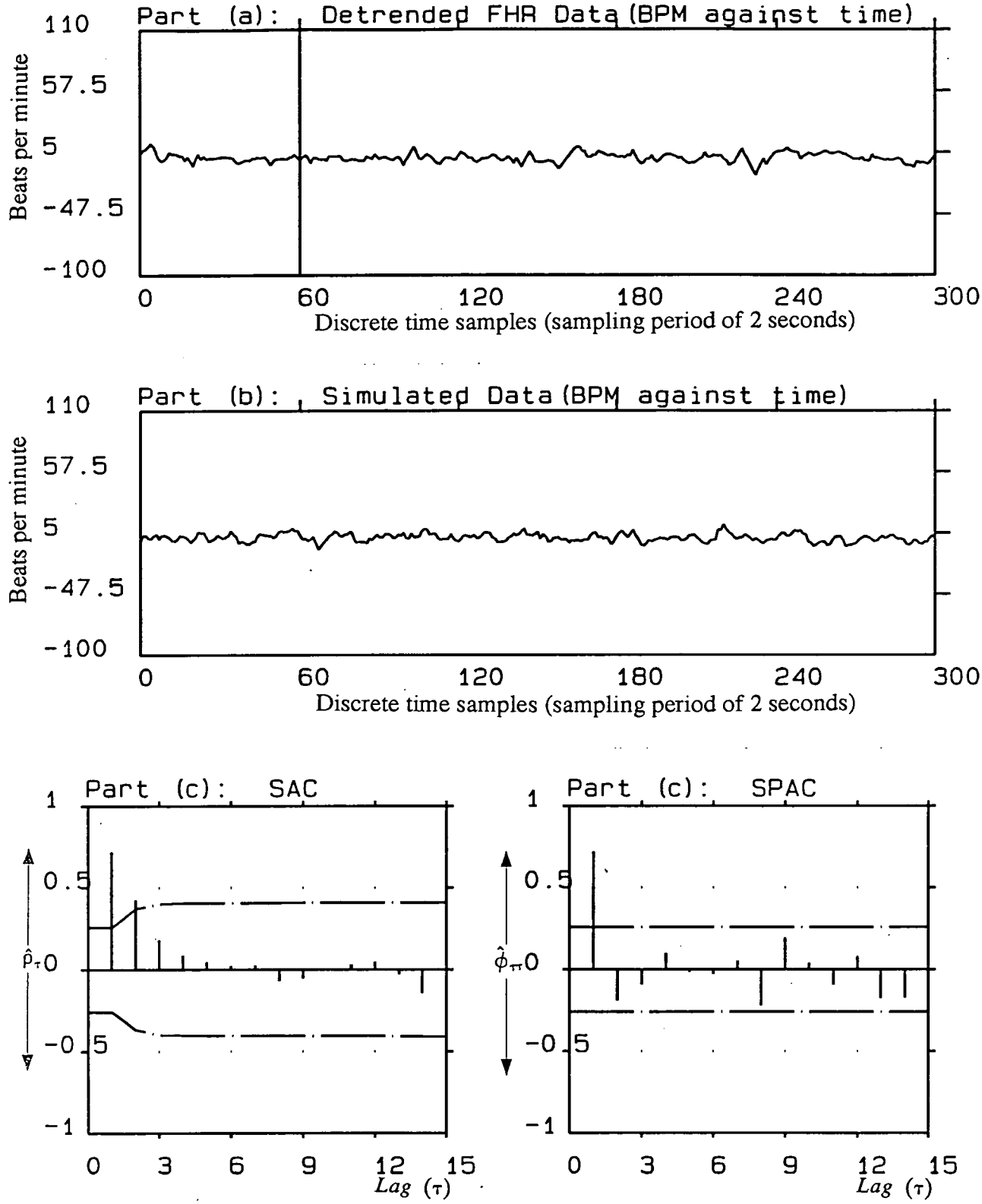
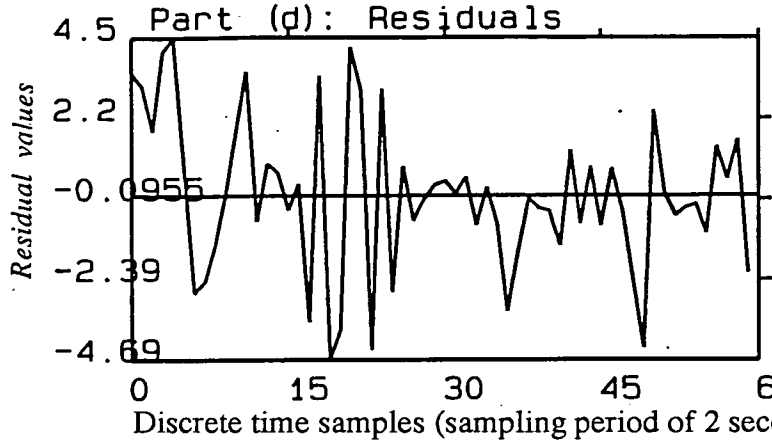


Figure 6.13: Part (a): Detrended averaged FHR (indirect phonocardiographic);
Part (b): Synthetic FHR data, generated from the parameters extracted from the boxed segment shown in part (a); the estimated parameter values were: $\hat{\phi}_1 = 0.9367$, $\hat{\phi}_2 = -0.2195$, $\hat{\sigma}_e = 2.1856$, $\hat{\sigma}_p = 3.4984$, $\hat{\mu}_{baseline} = 121.0168$; *Pseudo –periodicity* was not detected.
Part (c): SAC and SPAC of the first 60 synthetic data samples shown in part (b), indicating AR(1) correlation characteristics; The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.



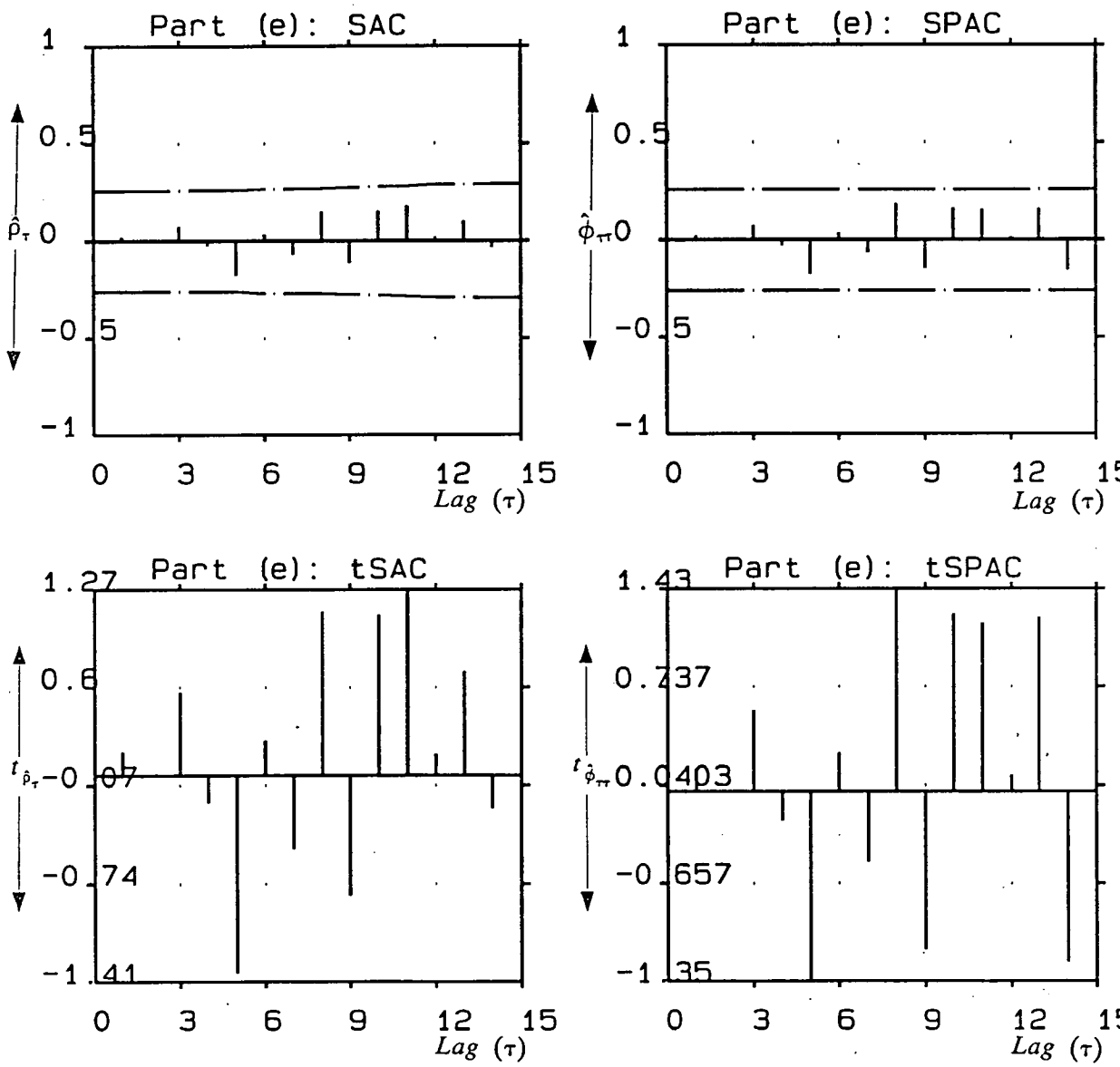
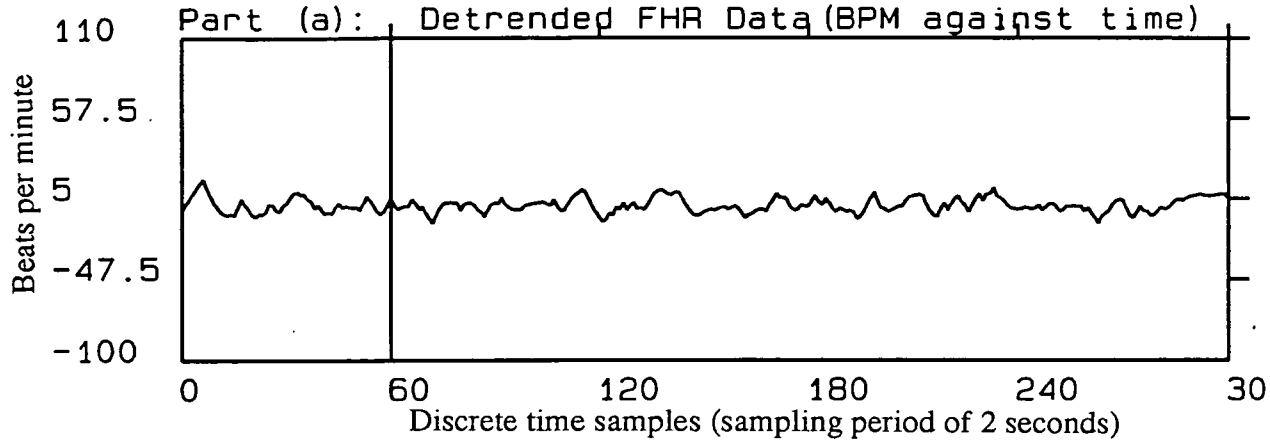
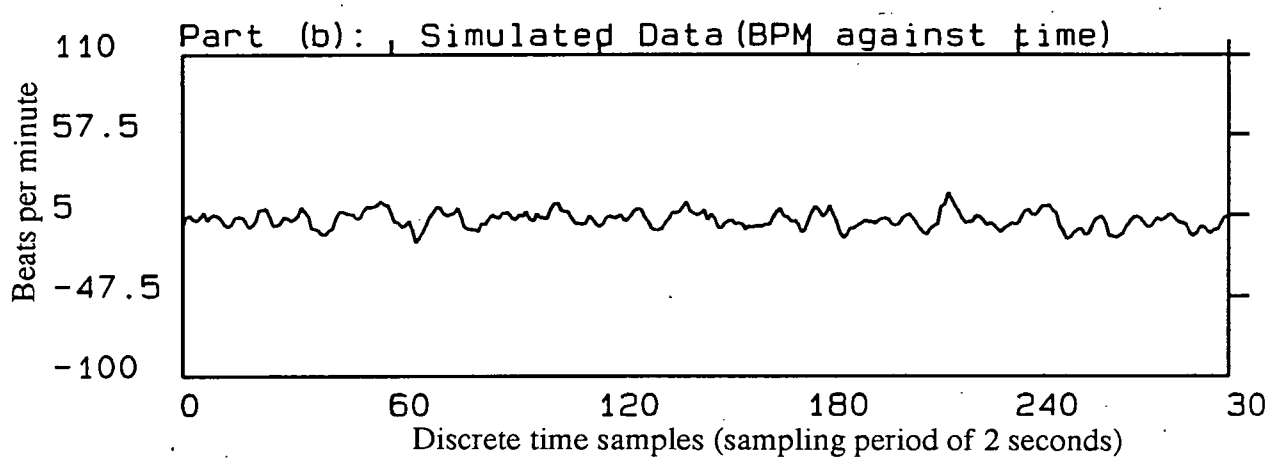


Figure 6.13: Part (d): Residuals (or Prediction errors or innovations) left after fitting an AR(2) model to the boxed segment shown in part (a).
Part (e): SAC, SPAC, t-statistics of SAC and SPAC of residuals shown above, indicating a white sequence; Q *–statistic* = 14.0317 $< \chi^2_{0.05}[13] = 22.3621$, further proof of the adequacy of the model fitted to the data.





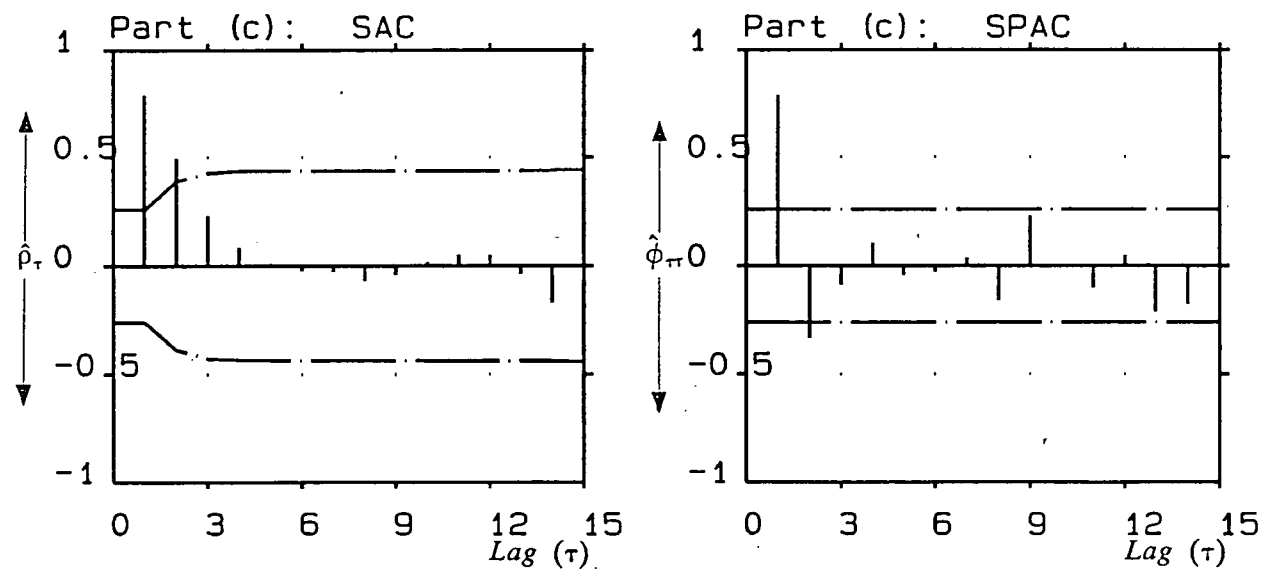
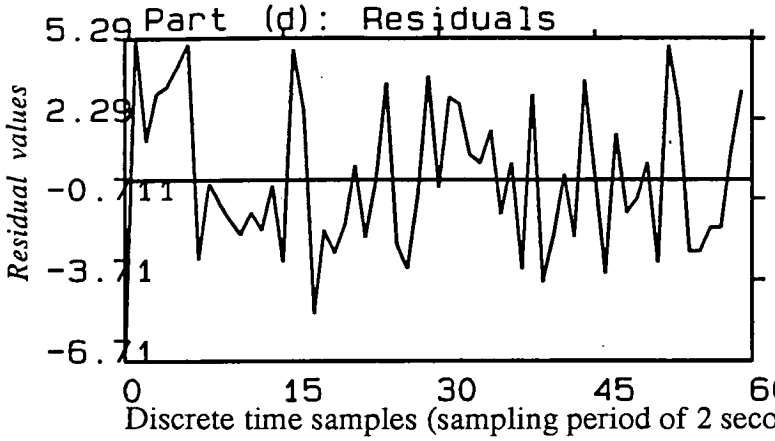


Figure 6.14: Part (a): Detrended averaged FHR (indirect phonocardiographic);
Part (b): Synthetic FHR data, generated from the parameters extracted from the boxed segment shown in part (a); the estimated parameter values were: $\hat{\phi}_1 = 1.1606$, $\hat{\phi}_2 = -0.3941$, $\hat{\sigma}_e = 2.7943$, $\hat{\sigma}_p = 5.4883$, $\hat{\mu}_{baseline} = 126.6974$; *Pseudo –periodicity* was not detected.
Part (c): SAC and SPAC of the first 60 synthetic data samples shown in part (b), indicating AR(2) correlation characteristics; The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.



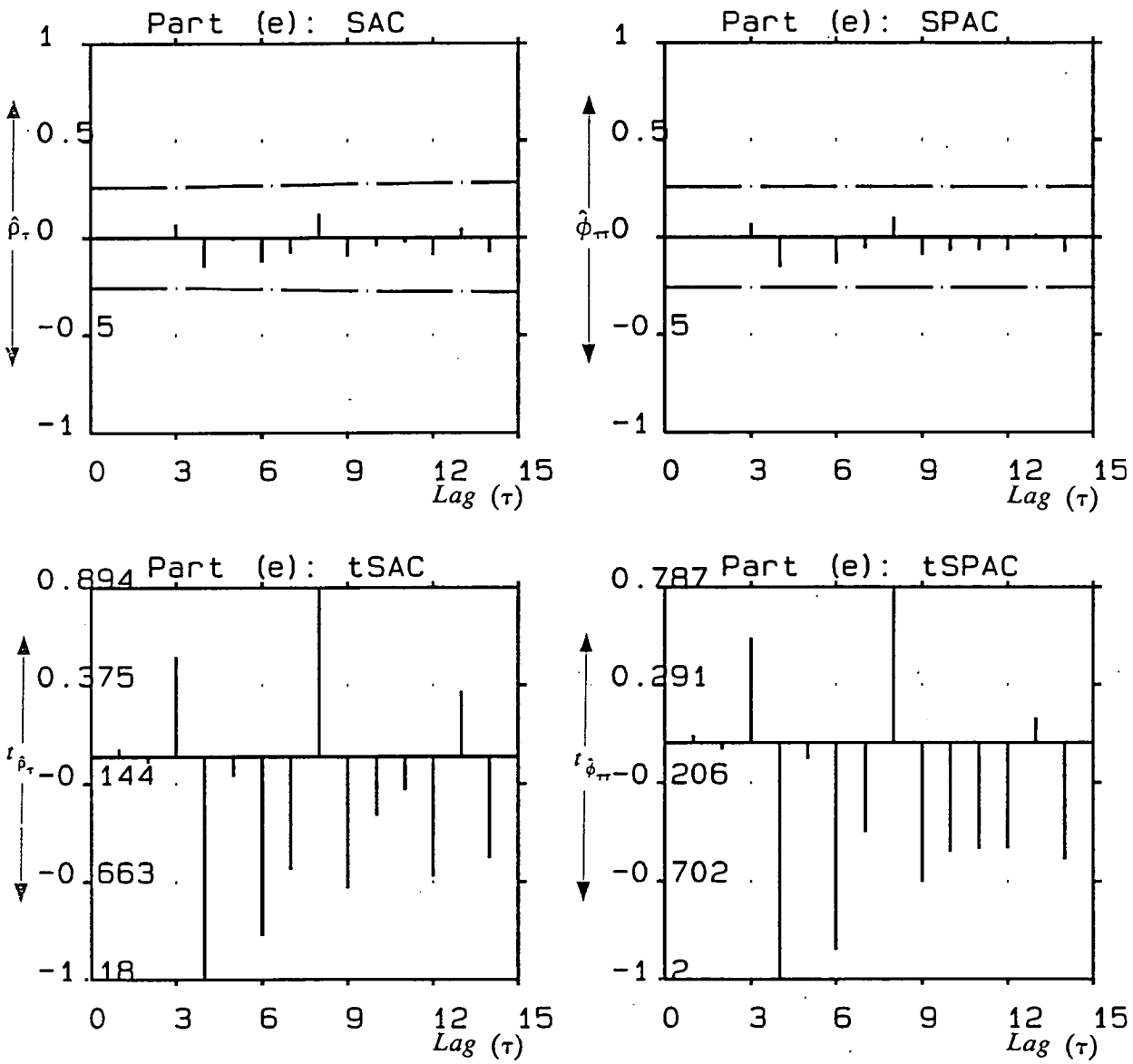


Figure 6.14: Part (d): Residuals (or Prediction errors or innovations) left after fitting an AR(2) model to the boxed segment shown in part (a).
Part (e): SAC, SPAC, t-statistics of SAC and SPAC of residuals shown above, indicating a white sequence; Q *–statistic* $= 7.2773 < \chi^2_{0.05}[13] = 22.3621$, further proof of the adequacy of the model fitted to the data.

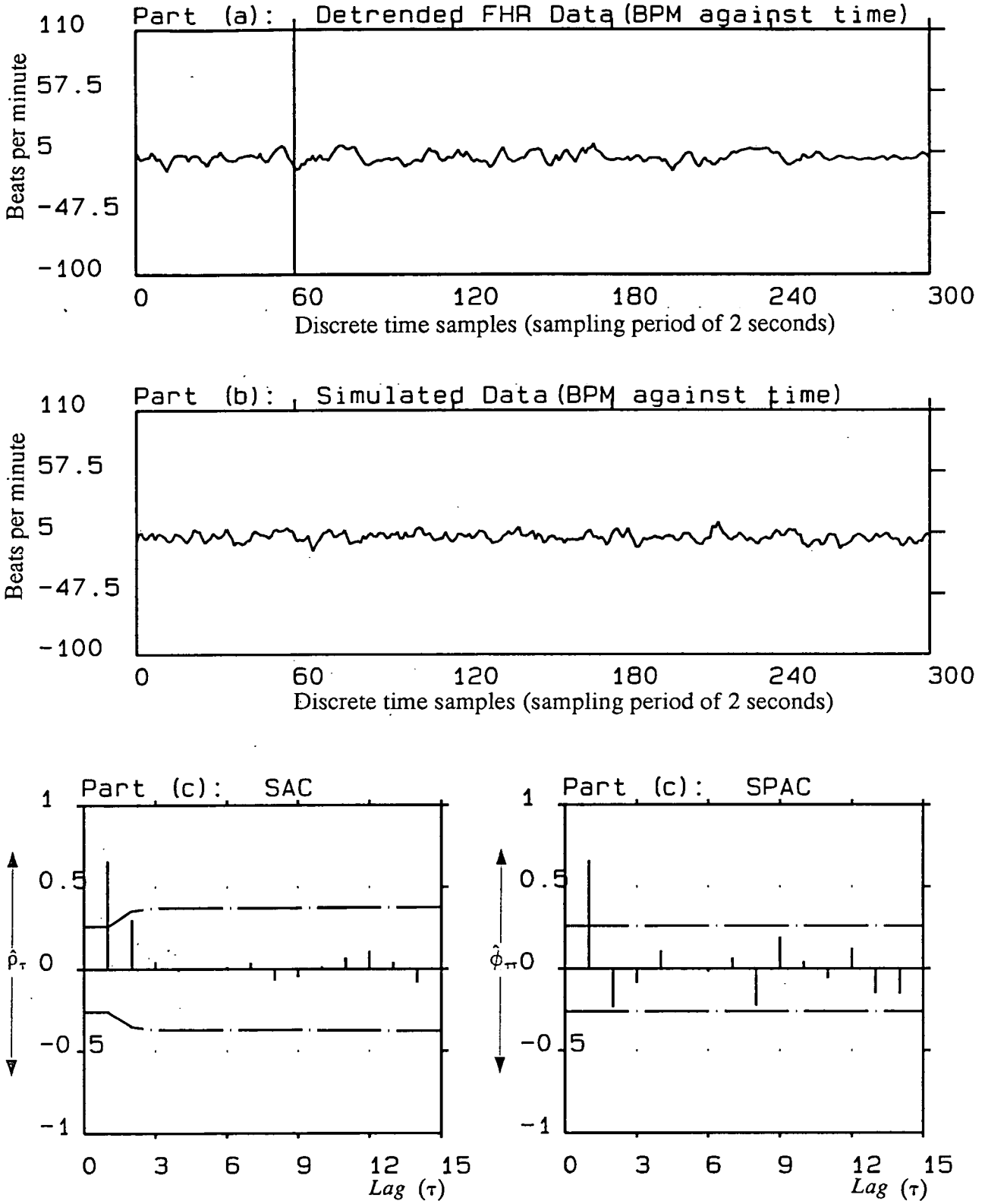
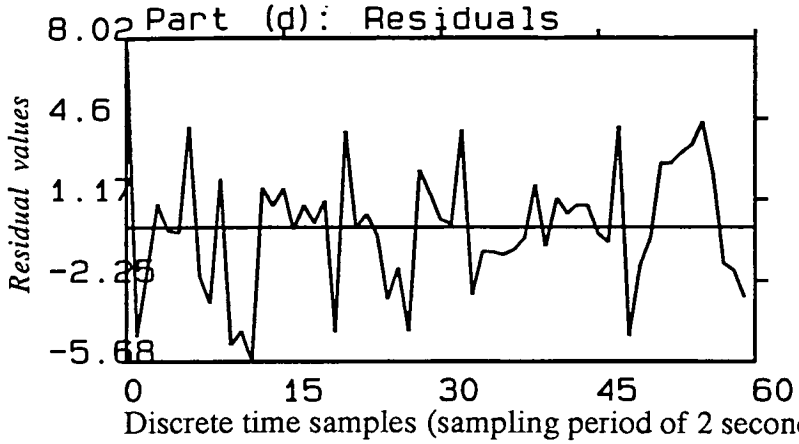


Figure 6.15: Part (a): Detrended averaged FHR (indirect phonocardiographic);
Part (b): Synthetic FHR data, generated from the parameters extracted from the boxed segment shown in part (a); the estimated parameter values were: $\hat{\phi}_1 = 0.8679$, $\hat{\phi}_2 = -0.2448$, $\hat{\sigma}_e = 2.7160$, $\hat{\sigma}_p = 3.9078$, $\hat{\mu}_{baseline} = 122.7360$; *Pseudo –periodicity* was not detected.
Part (c): SAC and SPAC of the first 60 synthetic data samples shown in part (b), indicating AR(1) correlation characteristics; The dot-dashed lines are confidence limits $\pm 2S_{\hat{\rho}_\tau}$ and $\pm 2S_{\hat{\phi}_{\tau\tau}}$ for SAC and SPAC respectively.



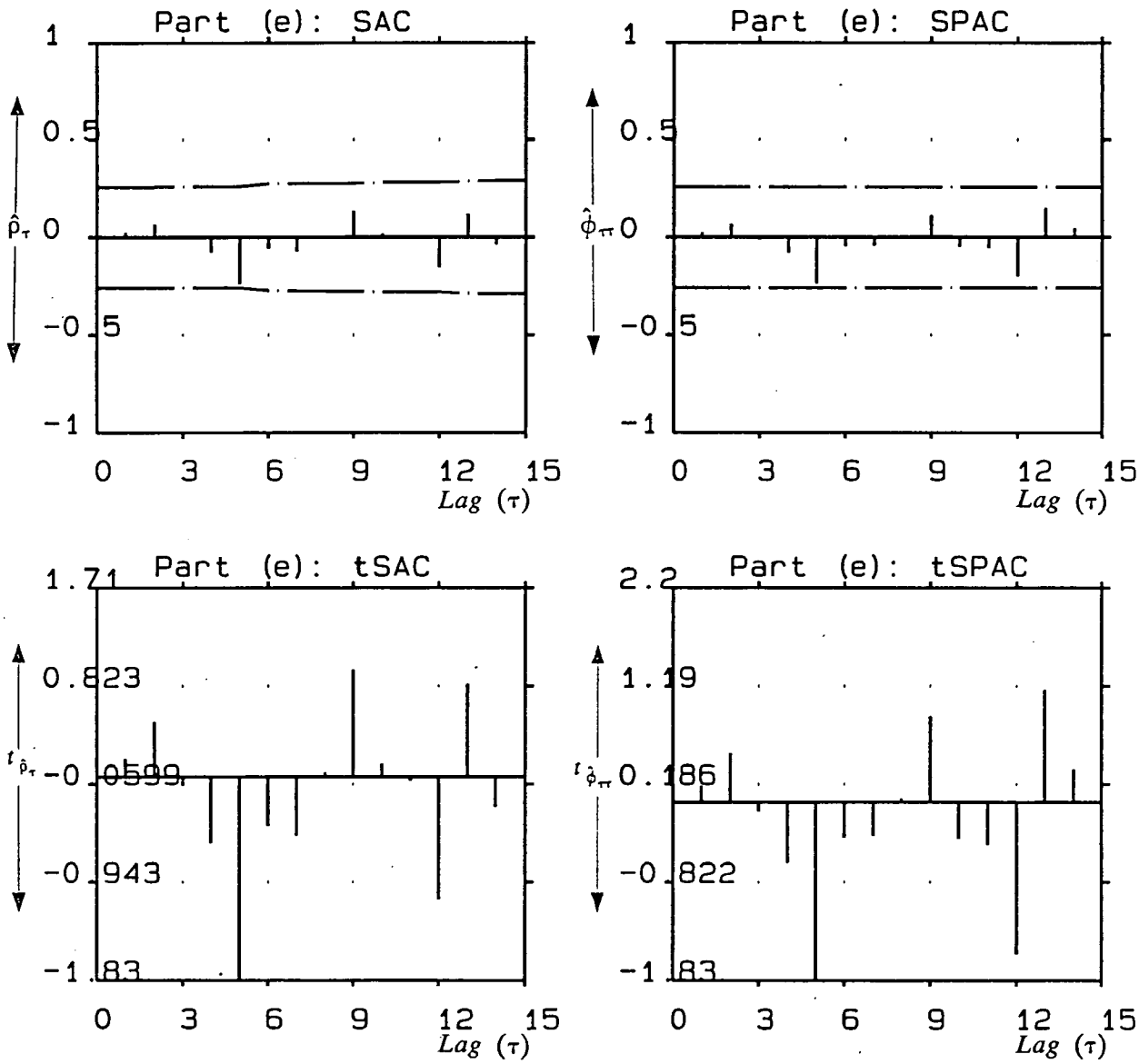


Figure 6.15: Part (d): Residuals (or Prediction errors or innovations) left after fitting an AR(2) model to the boxed segment shown in part (a).
Part (e): SAC, SPAC, t-statistics of SAC and SPAC of residuals shown above, indicating a white sequence; $Q-statistic = 14.6872 < \chi^2_{0.05}[13] = 22.3621$, further proof of the adequacy of the model fitted to the data.

Finally, as a test example, a two hour long, detrended, phonocardiographic FHR recording was segmented on a contiguous block basis, to 60 two-minute data blocks. These blocks were fitted to the AR(2) structure. All 60 fits had Q-statistics less than the critical value $\chi^2_{0.05}[13] = 22.3621$, further indicating the adequacy of AR(2) structure in terms of being able to parsimoniously model the underlying stochastic mechanism responsible for generation of long term FHR variability. The estimated parameters, together with the estimated standard deviation of each two minute data segment, and their corresponding Q-statistics are shown in table 6.3.

In summary, it has been shown in these examples that short data windows (detrended averaged FHR) of two minute duration (60 samples) are adequately represented in terms of five parameters. The first three are the ML estimated parameters $\hat{\phi}_1$, $\hat{\phi}_2$, and $\hat{\sigma}_e$ of the AR(2) model. These reflect the dynamic behaviour of long term FHR variability patterns. The other two are scalar quantities, $\hat{\mu}_{baseline}$ and $\hat{\sigma}_p$. The former represents the medically valuable information on FHR baseline level (mean estimated from the estimated FHR trend data) and the latter represents the genuine statistical spread of FHR variability over the estimated trend or baseline.

6.7 Stability and Detection of Pseudo-Periodicity

For an AR(2) or ARMA(2,q) process, the stationarity or stability requirement derived from the solution of the AR characteristic polynomial is a set of three conditions[59, 60, 62]:

$$\begin{aligned} &|\phi_2| < 1, \\ &\phi_2 + \phi_1 < 1, \\ &\phi_2 - \phi_1 < 1. \end{aligned} \tag{6.11}$$

These can be used to define a triangular stability region, as shown in figure 6.16. All

$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\sigma}_e$	$\hat{\sigma}_p$	$\hat{\mu}_{baseline}$	$Q-statistics$
1.0186	-0.4903	2.9846	4.6911	133.6858	9.0502
0.7955	-0.1769	2.6148	3.6048	129.9779	7.2018
0.8322	0.0324	2.5324	4.9662	136.5317	8.3812
1.1357	-0.3075	2.1540	4.5679	135.6231	6.2943
0.8388	-0.1720	2.6789	3.8941	137.1176	9.7468
0.8688	-0.1230	3.2299	5.1370	132.9440	9.8650
0.6640	-0.1455	2.8450	3.5291	129.8169	17.2637
1.2199	-0.3999	2.4207	5.3833	130.5223	16.0009
1.1596	-0.3807	2.3967	4.7749	126.4642	20.5814
0.8256	-0.2309	2.1329	2.9556	119.9582	12.2222
0.5923	-0.0780	2.5936	3.1137	117.7860	13.0775
0.9631	-0.1739	2.7523	4.8886	122.3603	14.1535
0.8590	-0.2245	2.5605	3.6871	117.7232	11.5534
1.3072	-0.5155	2.0834	4.8052	117.4582	10.9099
0.6755	-0.2837	1.9854	2.4348	116.8661	10.0651
0.9367	-0.2195	2.1856	3.4984	121.0168	14.0317
0.8120	-0.4181	1.8825	2.5277	122.9476	14.5637
1.1086	-0.3484	2.1012	3.9376	129.5325	14.7504
1.0009	-0.3544	2.5135	3.9897	136.1456	13.6528
0.8721	0.0006	1.9303	3.9521	138.5399	8.3700
1.2298	-0.3362	2.2821	6.1968	136.2210	5.8234
0.8524	-0.2420	1.6166	2.2908	126.9182	10.7955
1.1830	-0.4663	2.3284	4.4544	124.7288	7.5430
1.0193	-0.2369	2.3329	4.2391	127.8784	15.7133
1.1606	-0.3941	2.7943	5.4883	126.6974	7.2773
0.8679	-0.2448	2.7160	3.9078	122.7358	14.6872
1.0923	-0.2689	2.4869	5.0739	124.7542	12.5177
1.0445	-0.3761	2.7697	4.5912	128.1476	19.1942
1.0650	-0.2557	2.3957	4.6782	124.9714	9.0032
1.0647	-0.3514	1.3889	2.4088	128.4835	8.8580
1.2651	-0.4832	2.2447	4.9113	125.7790	9.5262
1.3375	-0.5223	2.8164	6.9168	127.1922	13.7011
0.9463	-0.3284	3.0263	4.5654	131.7461	17.6315
0.9107	-0.3858	3.5282	5.0734	133.6470	15.3197
0.9811	-0.2935	3.5791	5.7449	137.5846	5.7342
1.1116	-0.3459	3.4397	6.5023	139.5032	9.2903
1.1182	-0.2886	3.1966	6.7173	134.2713	16.7853
0.8347	-0.1583	3.2136	4.6946	133.3120	17.4048
1.1213	-0.2863	3.1581	6.7272	129.5476	14.8602
1.4536	-0.6520	3.0531	8.4748	132.7756	8.6779
0.6199	0.1093	2.5930	3.6326	121.0873	12.3534
1.1237	-0.2981	3.1267	6.5431	120.5403	22.2503
0.7596	-0.0993	3.2626	4.5360	122.4470	15.3521
1.1956	-0.4357	2.9385	5.8973	127.3905	15.9541
0.9564	-0.3267	4.3334	6.6158	127.9991	12.2209
0.9498	-0.1678	3.5100	6.1193	129.2430	13.5209
1.1218	-0.3592	3.2648	6.1951	134.4466	4.7922
0.8405	-0.0507	3.5279	5.8869	132.0457	17.4723
1.1969	-0.3821	3.1441	6.8044	128.2267	12.2167
1.0844	-0.3470	3.0192	5.4267	121.9501	15.6057
1.0891	-0.3900	3.4518	6.0331	126.6921	10.1166
0.8914	-0.1630	3.3573	5.2978	126.3513	5.4247
0.9679	-0.3850	3.1450	4.7640	123.3177	10.2924
1.1823	-0.4970	2.9830	5.6042	125.7053	6.4981
1.0453	-0.2249	3.1510	6.2023	131.2902	9.5311
1.1007	-0.3361	2.4839	4.6521	133.8820	10.4951
0.7803	-0.1243	2.8797	4.0311	132.0989	14.8523
1.1364	-0.3317	2.9936	6.0868	134.6735	11.9666
1.0293	-0.3411	3.4330	5.6965	133.2533	7.7820
1.1733	-0.3938	2.8519	5.7476	132.2469	4.5172

Table 6.3: A test example, showing the parameters that are estimated from two minutes data blocks from a 2 hours long average FHR file. These include the ML estimated parameters $\hat{\phi}_1$, $\hat{\phi}_2$, $\hat{\sigma}_e$ reflecting the dynamic of the FHR variability patterns and the scalar statistics $\hat{\mu}_{baseline}$ and $\hat{\sigma}_p$ reflecting the general FHR baseline level and the genuine FHR statistical spread over the baseline in each data window under analysis. Also the computed $Q-statistics$ for each set of estimated parameters are quoted, to test and observe the adequacy of the data fit and thus the model.

three conditions must be satisfied for an AR(2) or ARMA(2,q) model to be stationary.

In practice the conditions in equations (6.11) are applied to the estimates of ϕ_1 and ϕ_2($\hat{\phi}_1$ and $\hat{\phi}_2$), obtained at the parametric estimation stage. If the point associated with $\hat{\phi}_1$ and $\hat{\phi}_2$, falls within the stability triangle, the model will be stable. Otherwise this would be a sign that the data window fitted to the model is mean non-stationary.

This stability check is also a useful test for the nonlinear optimization routine, making sure that it operates and converges to a stable parametric point.

The stationarity conditions become complicated when $p > 2$. Fortunately ARMA models with $p > 2$, do not occur often in practice. When p exceeds 2, we can at least check the following necessary (but not sufficient) stationarity condition:

$$\phi_1 + \phi_2 + ... + \phi_p < 1. \tag{6.12}$$

A further valuable partitioning of the stability triangle can also be seen in figure 6.16. This partitioning divides the stability region into four sections, and has been achieved by analyzing the spectral behaviour of second order AR processes. In general, a discrete second order AR noise driven structure namely:

$$z(k) = \phi_1 z(k-1) + \phi_2 z(k-1) + e(k), \tag{6.13}$$

has the frequency response function:

$$H(f) = \frac{1}{1 - \phi_1 exp(-j2\pi fT) - \phi_2 exp(-j4\pi fT)}, \quad -\frac{1}{2T} \leq f < \frac{1}{2T}, \tag{6.14}$$

and hence the power spectrum

$$|H(f)|^2 = \frac{T\sigma_e^2}{1 + \phi_1^2 + \phi_2^2 - 2\phi_1(1-\phi_2)cos(2\pi fT) - 2\phi_2 cos(4\pi fT)}, \tag{6.15}$$

where σ_e^2 is the variance of AR input white driving noise, and T is the sampling period. For certain values of the parameters ϕ_1 and ϕ_2, (6.15) can give rise to a low-frequency or high-frequency spectrum, like the first order AR process. In addition to these spectra, it is possible to obtain spectra with a *peak* or a *trough* at an intermediate frequency f_o. This occurs when the following essential condition has been met[42, 105] :

$$|\phi_1(1 - \phi_2)| < |4\phi_2|. \tag{6.16}$$

This condition actually has also defined the spectral partitioning lines drawn within the stability triangle shown in figure 6.16. The frequency f_o at which the peak or trough occurs is given by:

$$f_o = \frac{1}{2\pi T}\cos^{-1}\left(\frac{-\phi_1(1 - \phi_2)}{4\phi_2}\right). \tag{6.17}$$

The detection of a peak, for a second order stochastic AR structure, is associated with the detection of *pseudo −periodicity* in the time series under observation[42]. This type of periodic behaviour is observed in detrended FHR time series which exhibit distorted oscillatory patterns.

The random process $e(k)$ which drives the model, causes random *phase −shifts* in the process which accounts for the *disturbed periodic* behaviour observed in time series such as detrended FHR. In some of the previous fifteen examples; when the condition (6.16) was met, pseudo-periodic behaviour was detected in the block under observation. In this case the estimated period, t_o $(=1/f_o)$, was also quoted along with each example.

This random periodicity, which for a long time was visually detected and has been a subject of discussion amongst the clinicians, can now be detected via this numerical

procedure. Using this routine has shown that the period of the cycles which was detected from detrended averaged phonocardiographic data, stretched from 10 seconds right up to 60 seconds.

Figure 6.17 shows 236 MLE estimates from 236 two minute data segments represented as points $(\hat{\phi}_1,\hat{\phi}_2)$, all falling within the stability triangle. They occupy the area of the stability triangle associated with low-frequency spectra, and peak frequency spectra. In a way this can be looked upon as a form of classifying or grouping the dynamic behavioural patterns observed in long term FHR variability which is represented by the detrended averaged FHR time series.

Although it was not the objective of this analysis to reconstruct exactly the detrended FHR time series, from the 236 abrupt block by block parametric estimates of the AR(2) parameters, the corresponding synthetic versions, were simulated and joined together to form the 8 hour long simulated data file shown in figure 6.18. Visually the AR(2) model based on contiguous blocks, has managed to track the general level of the variability observed in the real detrended 8 hour FHR file. Of the 236 data fits based on 2 minute contiguous blocks of the (approximately) 8 hour detrended FHR file (Fig. 6.18), 97 % were successful in terms of passing the *Ljung –Box Portmanteau* diagnostic lack of fit test, which further indicates the second order autoregressive structure as being an adequate model and a parsimonious representative of the stochastic mechanism responsible for the generation of the observed long term FHR variability patterns. Finally, it is also interesting to observe that the overall standard deviation statistic computed from the simulation very closely matches the one computed from the original detrended data (see Fig. 6.18 for these computed values).

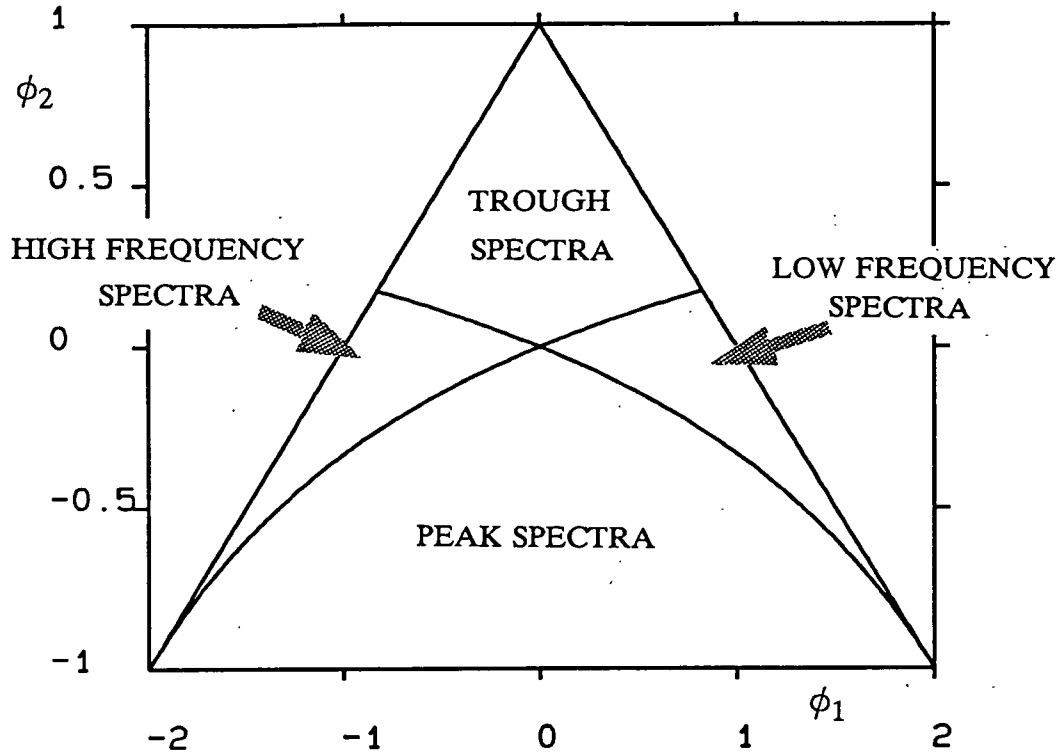


Figure 6.16: Stability Triangle (region) and Classification of Spectra for Discrete Second-Order Autoregressive Processes.

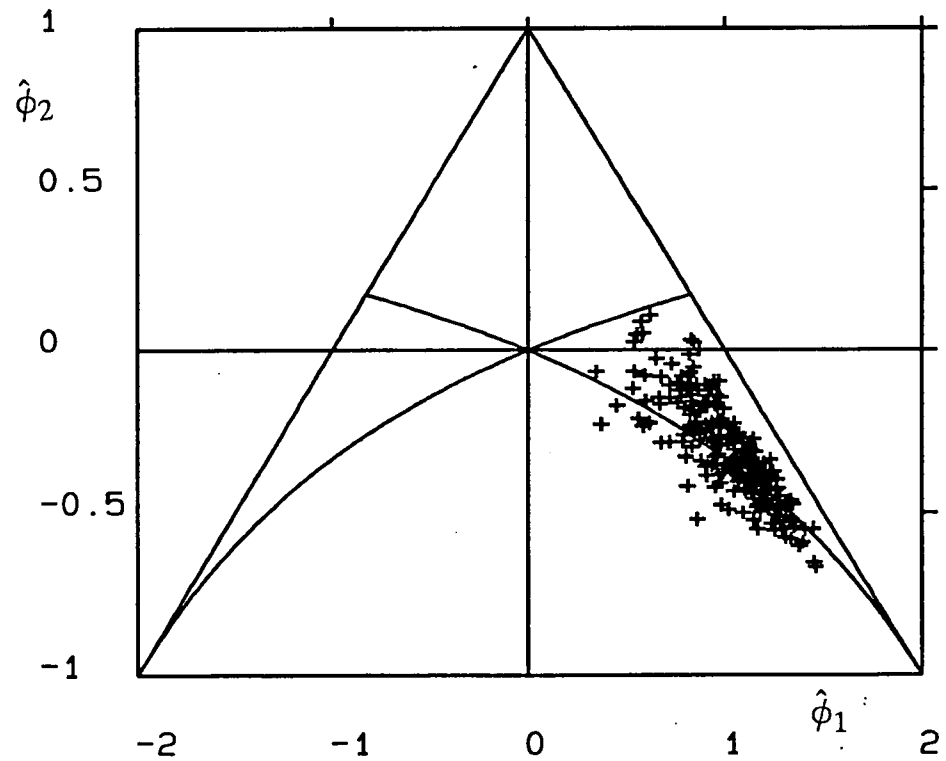
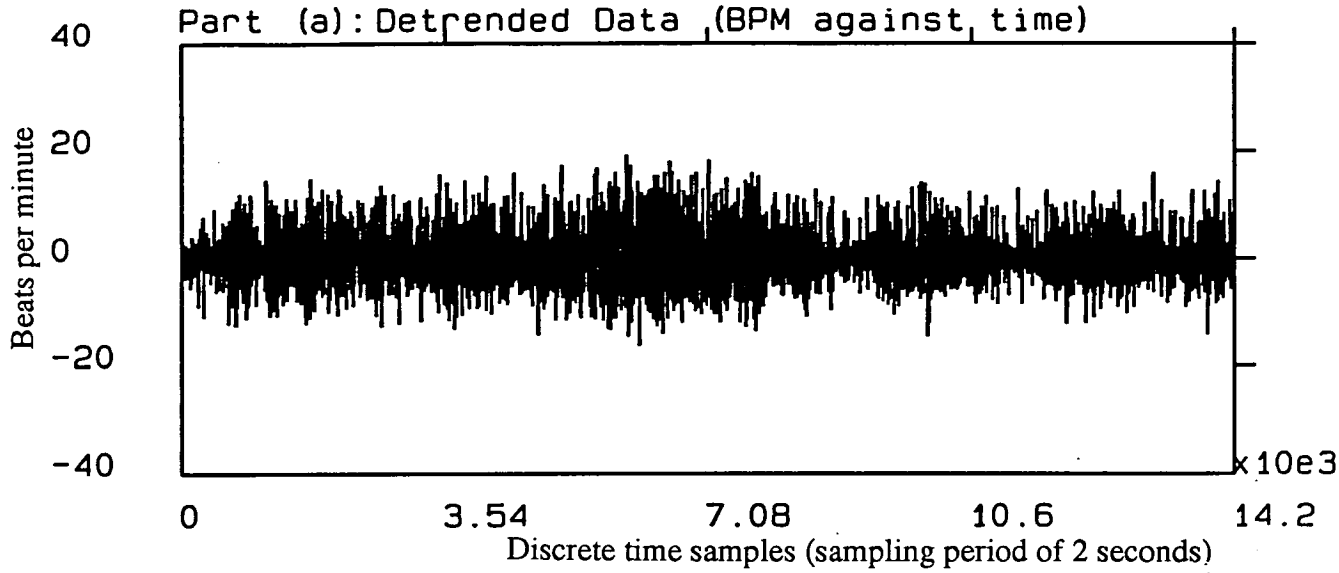


Figure 6.17: Speckle Pattern in the stability triangle derived from 236 points $(\hat{\phi}_1,\hat{\phi}_2)$, estimated from roughly 236 two minute different segments derived from an eight hour detrended averaged FHR recording.



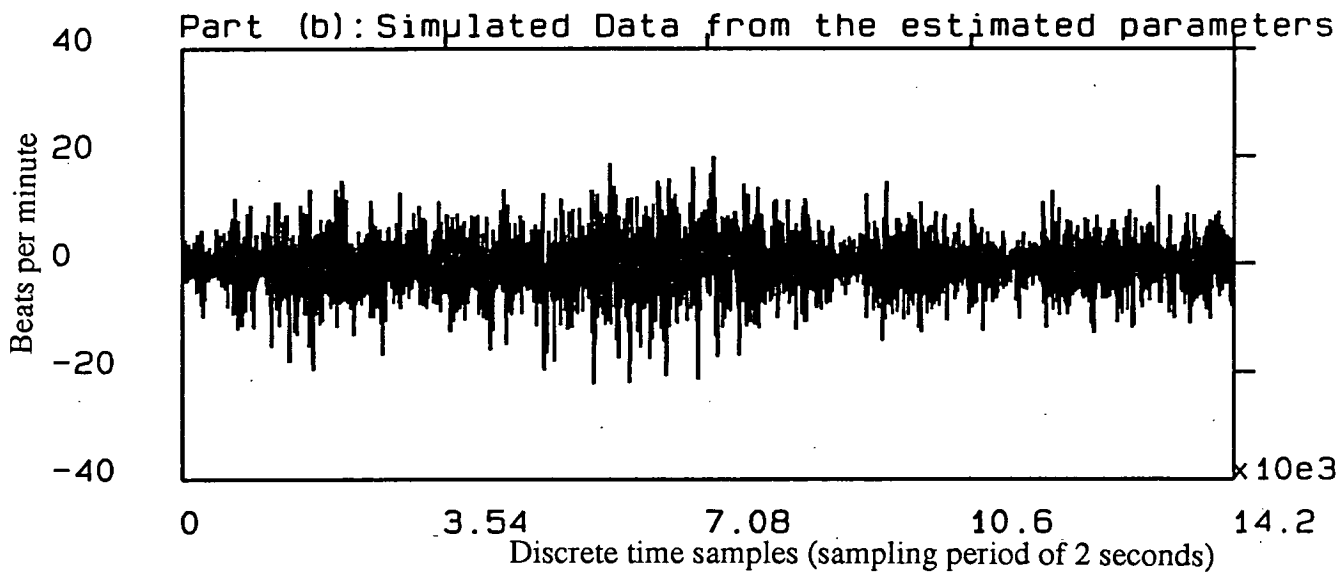


Figure 6.18: Part (a): Showing a complete eight hour detrended averaged FHR file. The estimated standard deviation (σ) for whole of this detrended FHR file was: "4.1991".
Part (b): Showing the corresponding (to Part a) simulated version produced from estimated parameters extracted from 2 minute quasi-stationary data windows on a contiguous basis. The estimated standard deviation (σ) for whole of this simulated file shown in part (b) was: "4.2381".

6.8 Conclusions and Discussion

The connection between the maximum likelihood estimation theory and Kalman filtering in defining the full log-likelihood function for a dynamic system was explained.

The non-linear log-likelihood criterion function was successfully optimized, through a proper reparameterization approach introduced in this chapter for constraining the area of search in the parametric space, and the use of a multidimensional non-gradient technique, namely the direction set Powell's method. The optimization routine never diverged and was stable, producing parameter estimates which fell within the stable boundaries of AR(2) parametric space.

Overall, the long term FHR variability represented by the detrended, two second averaged, phonocardiographic, FHR estimates, is stochastically well-fitted to a univariate stochastic AR(2) process. This is a second order difference structure, which as part of its inherent features, can exhibit pseudo-periodic oscillatory behaviour under suitable conditions. This type of random periodic oscillation can be easily seen when visually analyzing the detrended FHR time series. The periodicity, that was detected in the random variability patterns, stretched roughly from 10 to 60 seconds.

The parameters estimated are a good representation of the dynamics of long term FHR variability (detrended data). The synthetic data generated from the estimated parameters, bear a close visual resemblance to the original, detrended, FHR data, providing additional confirmation that the long term variability data are adequately represented.

By contrast, other scalar one-dimensional measures of long term variability cannot distinguish among the variety of statistically correlated patterns which are seen clinically.

Thus the AR-parameters could be used to discriminate between variability patterns observed in different short locally stationary windows of observations. The technique can also be looked upon as means of data reduction. This could be of value when dealing with a highly redundant information source, such as long term recordings of FHR.

By using the partitioned stability triangle introduced for an AR(2) process, it was clearly observed from ML parameters estimated from an eight hour detrended FHR recording (averaged phoncardiographic), that the series mostly either exhibits low-frequency type spectra or peak-frequency type spectra. By observations of the values of the estimated AR(2) parameters, and the relative position that they occupy in the stability triangle, we can conclusively detect if there has been any pseudo-periodic behaviour during the period of observation.

The type of speckle patterns observed (like the one in Fig. 6.17), can be looked upon as a way of summarizing the dynamics of random (correlated) variability patterns observed in detrended FHR recordings of long duration. This could have considerable medical potential for the early detection of problems. Perhaps, the AR(2) stability triangle speckle pattern could be correlated to medical conditions.

CHAPTER 7

CONCLUSIONS AND PROPOSALS FOR FURTHER WORK

7.1 General Summary, Discussion and Conclusions

The general aim of this project was to find suitable, comprehensive and meaningful numerical techniques for the purpose of objective quantitative analysis of patterns observed in FHR time series obtained from the period of pregnancy before labour (the antepartum period). In order to best approach this problem however, the FHR components of interest to the obstetricians and midwifes had to be understood; information that was not coherently and readily available.

Essentially the FHR components of interest during the antepartum period are the FHR baseline level, FHR accelerations, and finally the long term FHR variability (FHRV) which generally represents correlated fluctuations over the baseline. In current routine clinical practice, these components are still assessed visually from graphical records of FHR. In this project all of these components have been analyzed numerically in a comprehensive manner.

An important observation has been made concerning medical experts evaluating FHR patterns. Their assessment seems to be *binary* and *qualitative*, that is what they are looking for is, either present or not present, and results are very much dependent on

how good have they become in visually interpreting FHR time series.

In regards to long term FHR variability, no meaningful, conclusive, medical interpretation has been made concerning the variety of random correlated periodic patterns observed. Generally, what the medical experts want to see (as a positive sign on fetal condition), is a good level of FHR random variation over the baseline, preferably above 5 BPM (peak to peak) and less than 25 BPM.

In this study we were not in a position to make medical interpretations of the long term FHR variability patterns. However, we attempted to find an efficient numerical representation of these variety of correlated random periodic variability patterns. Hopefully, this will be of use to the medical community.

Prior to the analysis of long term variability (variation over the baseline) or the detection of accelerations, proper estimation and removal of the FHR baseline (to give detrended data) was an essential.

In general there were two methods of estimating and removing the slow underlying rhythms (baseline or trends) observed in FHR time series. One way was to fit to the data, on a contiguous block basis, different deterministic models such as the simple linear trend model or nonlinear polynomial models. However, when dealing with long hours of measurements, this approach for over-all estimation of and removal of the trend will result in unquantifiable statistical bias in the estimated baseline and in the detrended data which is used for long term variability analysis. This problem was remedied by use of some form of digital filtering technique, namely the first order linear bi-directional autoregressive filter. Because of its bi-directional structure, phase shift was completely eliminated.

As result of being able to obtain an unbiased estimate of the FHR baseline, the detri-

mental effect of baseline (mean nonstationarity) on the scalar standard deviation statistic which is the general statistic used by the medics for their quantitative analysis of FHR variability was shown. It was concluded that in order to represent the genuine spread of the variability over the baseline, through the use of the scalar standard deviation statistic (or indeed any related statistic), it is vitally important to use detrended FHR data instead of the original.

The detrended FHR was still nonstationary in the second moment (variance), i.e. its variance continuously changed with time. Therefore in order to capture correctly the development with time of long term FHR variability, the shortest possible observational epoch should be used for classical scalar statistical analysis.

In our analysis the detrended FHR was segmented into equally sized one minute duration data blocks and their standard deviation taken. The length of the windows were short enough to have local statistical stationarity and long enough to be considered as containing a large sample, for the purpose of having unbiased standard deviation estimation. This approach can be regarded upon as a one-dimensional statistical way of representing quantitatively, evolutionary variance (spread) of long term FHR variability.

The next step taken, was to compute the relative frequency histogram of these standard deviation statistics. This resulted in a useful tool to show the general status of the FHR variability over a long term antepartum FHR recording in an efficient and highly informative manner. These histogram shapes can be further summarized using moment based statistics, namely the mean, variance, the coefficient of skewness and the coefficients of kurtosis. This form of histogram analysis will be a valuable tool, when hundreds of long duration recordings are being analyzed.

The other important component of interest, namely FHR accelerations also benefited from the unbiased estimation of FHR baseline. A rule based accelerations detection routine was devised, based on the rate of change of the estimated baseline, the original FHR time series, and the rule of what constitutes a genuine acceleration. The procedure scanned two data sets (the original FHR, and its estimated rate of change of baseline), for a possible acceleration; checked to see if it was genuine, and if so, then proceeded to scan back and forth in the vicinity of the detected acceleration, in order to extract relevant numerical information. The routine is non-exhaustive unlike the human expert and can accurately scan hours of antepartum recorded FHR data for detection of accelerations.

Although the classical statistical technique properly analyzed the evolutionary variance of variability over the baseline, it lacked the ability to actually represent the variety of correlated (sometimes pseudo-periodic) patterns observed in long term FHR variability. Such scalar statistics were useless if numerical discrimination between different types of non-white variability patterns was required. Clearly this was another very important problem which needed addressing.

To tackle this problem there were two important objectives which had to be adhered to. These were simplicity and ease of understandability of the procedure that would be used to quantitatively represent long term FHR variability patterns. This was essential if the statistics were to find application by the medical community. Also this might allow physiological significance to be attached to them once extensively applied to large volumes of data.

The available data (i.e. the detrended averaged phonocardiographic FHR) was discrete and stochastic in nature. Thus the model that could represent the correlated dynamic patterns observed in long term variability had to be a stochastic model. As a result this led to application of stochastic univariate time series analysis. Hopefully

the relevant theorems along with the results that have been presented in the previous chapters are coherent and understandable.

Extensive application was made of time series analysis steps, namely identification (or specification), model fitting (or parametric estimation) and diagnostics; which was explained thoroughly in the previous two chapters. By this means, the adequacy of an AR(2) model for modeling short, locally stationary, segments of long term FHR variability was proven, and its adequacy as a parsimonious model for long term variability could not be rejected.

The synthetic data generated from the parameter estimates, had a close visual resemblance to the original detrended data (representing long term variability). By contrast other scalar measures like standard deviation could not distinguish among the variety of observed variability patterns. However, the parameters of the selected model could be used to discriminate between the variety of correlated variability patterns observed. The technique could also be looked upon as a means of data reduction. This is of value, since long duration FHR recordings are required during the antepartum period.

Although long term FHR variability patterns have, in general, not yet been classified medically, and the medical meaning behind these patterns is not yet understood, the work of this project has led to the provision of meaningful numerical features which can be used to train and run a long term FHR variability pattern classification system. This could, for example, be based on a Neural Network approach (see figure 7.2). Although this original objective was not possible, the long term variability was classified in terms of its general spectral behaviour, via the use of a spectrally partitioned stability triangle defined for a second order difference structure introduced in the previous chapter.

The speckle pattern observed (see Fig. 6.17) from a normal eight hour recording, clearly fell into two spectral groups; low frequency and peak frequency. The ones which fell into the peak frequency section of the triangle represented data which exhibited pseudo-periodic behaviour.

The speckle pattern observed in the stability triangle, may be regarded as a way of summarizing the information about the dynamics of long term FHR Variability patterns. The general shape of speckle pattern may also change for different patients with different medical conditions. If this is the case, then such plots may prove to be a valuable tool for assessment of fetal condition from long term antepartum FHR recordings.

Figure 7.1 summarizes what valuable information the numerical analysis routines introduced in this thesis may provide to medical researchers, obstetricians and midwifes. This could lead, the first time, to accurate objective numerical screening of FHR, which should result in a more consistent and reliable interpretation of indirect FHR information during the antepartum period.

7.2 Proposals for Further Work

As was explained in the previous section, an FHR monitoring system based on long term variability patterns, will only be of use if these patterns have predictive or diagnostic value. On the assumption that they have, a medical long term variability pattern classification system such as the one shown graphically in figure 7.2 would be feasible. This would be an important area of work for the future, using numerical parametric features related to the patterns derived from the time series analysis approach that has been introduced in this thesis.

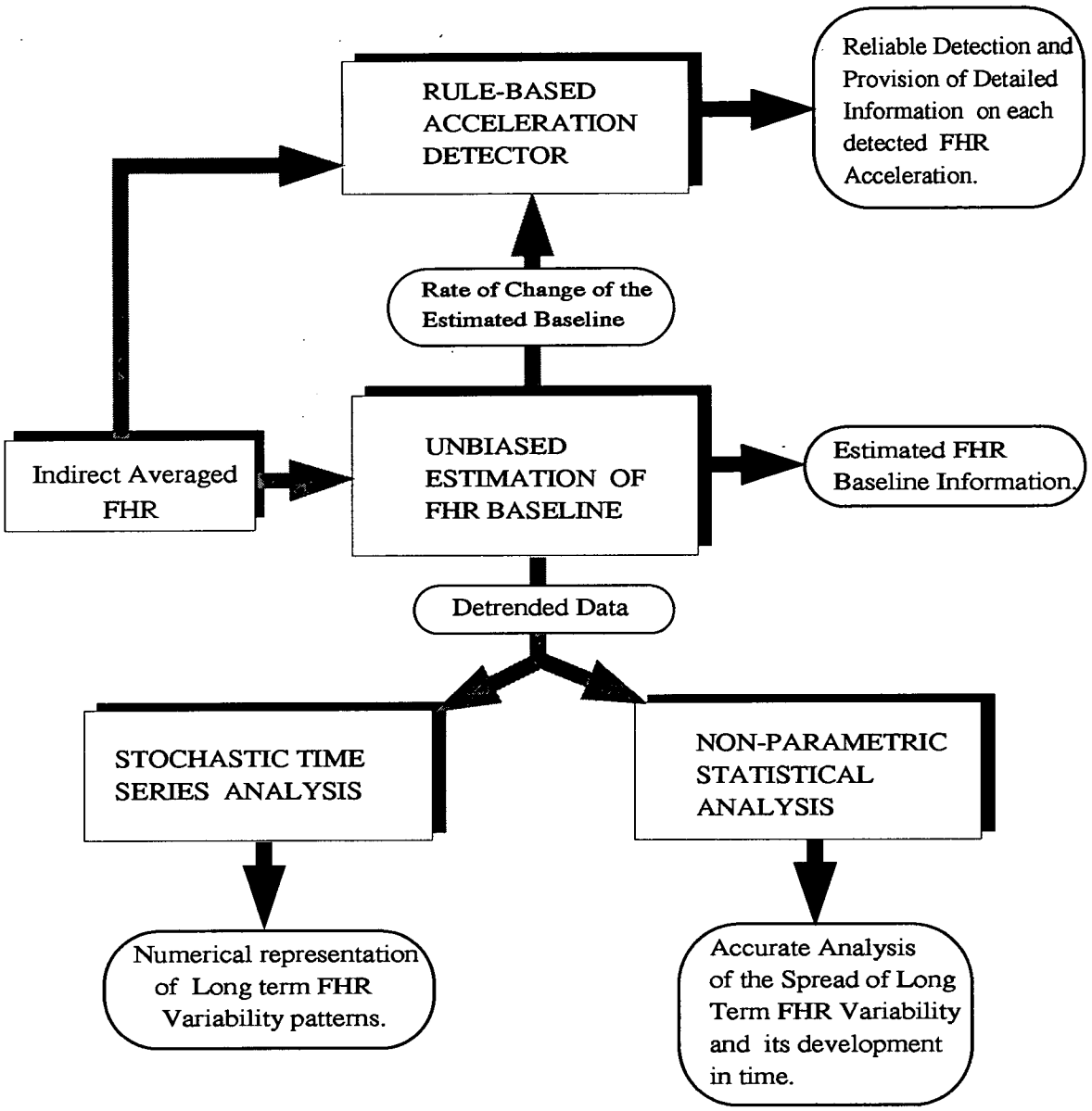


Figure 7.1: General Summary, showing the roots taken to provide detailed relevant numerical information on all of the antenatal FHR components of interest to doctors, facilitating the genuine objective numerical analysis of FHR time series.

Another topic for research would be analysis of the variability histogram shapes (chapter 4) from different patients with different medical conditions. The same applies for the speckle plots in the stability triangle introduced in the previous chapter. What would the shape of speckle distributions from recordings made from different patients with different symptoms reveal? These would be valuable findings, and are very much dependent on availability of a large volume of antepartum FHR data from large numbers of patients exhibiting a variety of known medical conditions.

Finally, the analysis and processing reported in this thesis, was performed offline, and the numerical analysis routines operated disjointly. For on-line use, e.g., a bedside monitor these routines should be adapted to operate in real time and they should be integrated to provide an objective numerical analyzer of the main FHR components of interest, as a prelude towards the final goal of developing a comprehensive fetal screening expert system (decision support tool), which will take into account other fetal parameters and medical information, such as fetal breathing rate, fetal movements, systolic time interval, and information such as medical condition of mother, fetal age, etc.

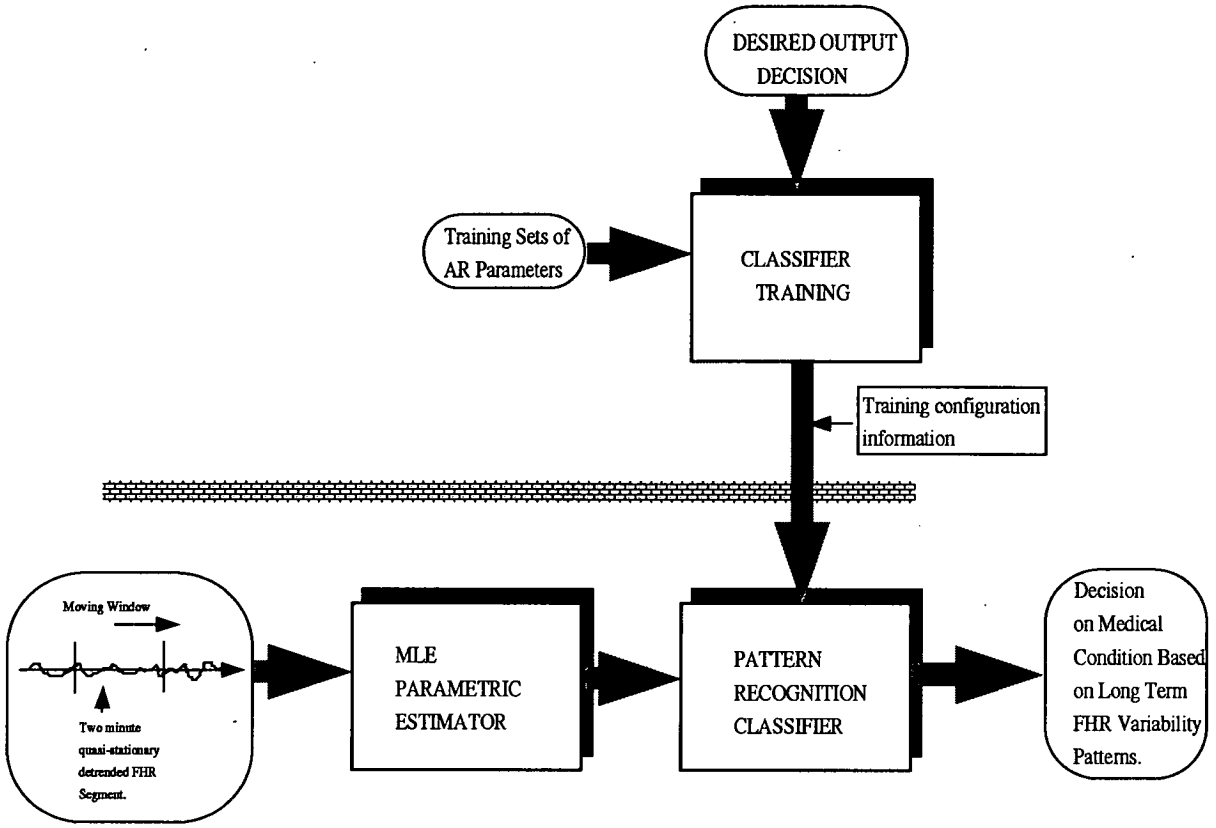


Figure 7.2: Time series-based system proposed for predicting fetal condition, from Long Term FHR variability patterns, the pattern recognition method will use the AR - parameters as features.

References

1. R.K. Freeman and T.J. Garite, *Fetal Heart Rate Monitoring*, Williams and Wilkins, 1981.

2. J. E. E. Fleming, *Fetal Monitoring*, pp. 332-339, Pergamon Press, Glasgow, UK, 1984. From "Physics in Medicine & biology Encyclopedia;T.F. McAinsh"

3. W. Kunzel, *Fetal Heart Rate Monitoring*, Springer-Verlag, Heidelberg, 1985.

4. M.C. Carter, "Fetal Monitoring," *J. Biomed. Eng.*, vol. Vol. 10, pp. 527-532, Butterworth & Co, 1988.

5. L. Cromwell, F.J. Weibell, and E.A. Pfeiffer, *Biomedical Instrumentation and Measurements*, Prentice-Hall, New Jersey, 1980.

6. L. Stryer, *Biochemistry*, San Francisco, 1981.

7. F.P. Zuspan, E.J. Quilligan, J.D. Iams, and H.P. vanGeijn, "NICHD Consensus Development Task Force Report. Predictors of intrapartum fetal distress. The role of fetal monitoring.," *Am. J. Obstet Gynecol*, vol. 135, pp. 287-290, 1979.

8. C. Sureau, "Fetal monitoring, present and future-Discussion," *Europ. J. Obstetrics and Gynaecology*, vol. 24, pp. 119-121, 1987.

9. J.A.D. Spencer, *Fetal Heart Rate Variability*, pp. 103-122, 1987. From "Progress in Obstetrics and Gynaecology"

10. E.H. Hon, *An atlas of fetal heart rate patterns*, Harty Press, New Haven, Connecticut, 1968.

11. T.H. Frank, O.R. Blaumanis, R.K. Gibbs, and R.K. Wells, "Adaptive Filtering in ECG Monitoring of the Fetal Heart Rate," *J. Electrocardiology*, vol. Vol 20.(Supplemental Issue), pp. 108-113, 1987.

12. D. Callaerts, B. DeMoor, J. Vandewalle, and W. Sansen, "Comparison of SVD methods to extract the fetal electrcardiogram from cutaneous electrode signals," *Medic. Biol. Eng. Comp.*, vol. 28, pp. 217-224, 1990.

13. T. Wheeler and A. Murrills, "Patterns of fetal heart rate during normal pregnancy," *Brit. J. Obstet. & Gynaec.*, vol. 85, pp. 18-27, 1978.

14. V. Kariniemi, T. Katila, H. Laine, and P. Ammala, "On-line quantification of fetal heart rate variability," *J. Perinatal Med.*, vol. 8, pp. 213-216, 1980.

15. M.C. Carter, P. Gunn, and R.W. Beard, "Fetal heart rate monitoring using the abdominal fetal electrocardiogram," *Brit. J. Obstet & Gynaec.*, vol. 87, pp. 396-401, 1980.

16. D.L. Tuck, "Improvement in Doppler Ultrasound human fetal heart rate records by signal correlation," *Med. Biol. Eng. Comput.*, vol. 20, pp. 357-360, 1982.

17. D.G. Talbert, J. Dewhurst, and D.P. Southall, "New Transducer for detecting fetal heart sounds: Use of compliance matching for maximum sound transfer," *Lancet*, vol. 1, pp. 426-27, 1984.

18. D.G. Talbert, W.L. Davies, F. Johnson, N. Abraham, N. Colley, and D.P. Southall, "Wide bandwidth fetal phoncardiography using a sensor matched to the compliance of the mother's abdominal wall," *IEEE Trans. Biomed. Eng.*, vol. 33, pp. 175-80, 1986.

19. M.N. Ansourian , "Digital Signal Processing for the Analysis of Fetal Breathing Movements ," Ph.D. Thesis, University of Edinburgh , 1989.

20. H.G. Goovaerts, O. Rompelman, and H.P.van. Geijn, "A Transducer for Detection of Fetal Breathing Movements," *IEEE Trans. Biomed. Eng.* , vol. 36, pp. 471-478, 1989 .

21. L.D. Devoe, "The Fetal Biophysical Profile: Antepartum Assessment Using a Programmed Microcomputer," *J. Clinicl. Eng.*, vol. 11, pp. 285-289, Quest & Co, Jul-Aug, 1986.

22. D.M. Sherer, M.N. Nawrocki , N.E. Peco, L.A. Metlay, and J.R. Woods, "The Occurrence of Simultaneous Fetal Heart Rate Accelerations in Twins During Non-

stress Testing," *Obstetrics and Gynecology*, vol. 76, pp. 817-821, 1990.

23. G.S. Dawes, G.H.A. Visser, J.D.S. Goodman, and D.H. Levine, "Numerical analysis of the human fetal heart rate: modulation by breathing and movement," *Am. J. Obstet. & Gynaec.*, vol. 140, pp. 535-544, 1981.

24. P.J.D. Wickham, G.S. Dawes, and R. Belcher, "Development of methods for quantitative analysis of the fetal heart rate," *J. Biomed. Eng.*, vol. 5, pp. 302-308, 1983.

25. G.S. Dawes, C.W.G. Redman, and J.H. Smith, "Improvements in the registration and analysis of fetal heart rate records at the bedside," *Br. J. Obstet. & Gynaec.*, vol. 92, pp. 317-325, 1985.

26. H.P. vanGeijn, "The value and interpretation of fetal heart rate patterns," *Perinatal Medicine*, pp. 71-78, H.T.P. Press Ltd., Lancaster, 1985.

27. J.G. Nijhuis, H.F. Prechtl, C.B. Martin, and R.S.G.M. Bots, "Are there behavioural states in the human fetus," *Early Human Dev.*, vol. 6, pp. 177-95, 1982.

28. J.G. Nijhuis, "Behavioural states in the human fetus," *thesis*, University of Nijmegen, Netherlands, 1984.

29. R. Natale, C. Nasello, and R. Turlink, "The relationship between movements and accelerations in fetal heart rate at twenty four to thirty two weeks gestation," *Am. J. Obstet. & Gynaec.*, vol. 148, pp. 591-595, 1984.

30. R. Caldeyro-Barcia, C. Mendez-Bauer, and J.J. Poseiro, *Control of the human fetal heart rate during labour.In: Cassels D.E(ed) The heart and circulation in the newborn and infant,* pp. 7-35, Grune & Stratton, New York, 1966.

31. K. Hammacher, K.A. Huter, J. Bokelmann, and P.H. Werners, *Fetal heart frequency and perinatal condition of the fetus and newborn*, 166, pp. 349-360, 1968.

32. R.W. Beard, G.M. Filshie, C.A. Knight, and G.M. Roberts, "The significance of the changes in the continuous fetal heart rate in the first stage of labour," *J. Obstet. & Gynaec. of the Br. Comm.*, vol. 78, pp. 865-881, 1971.

33. R.H. Paul, A.K. Suidan, S.Y. Yeh, B.S. Schifrin, and E.H. Hon, "Clinical fetal monitoring. VII. The evaluation and significance of intrapartum baseline FHR variability," *Am. J. Obstet. & Gynaec.*, vol. 123, pp. 206-210, 1975.

34. H.P. vanGeijn, "Fetal monitoring-present and future: The evaluation of fetal heart rate patterns," *Europ. Jnl. of Obst & Gyn.*, vol. 24, pp. 117-119, 1987.

35. F.A. Manning, L.D. Platt, and L. Sipos, "Antepartum fetal evaluation: Development of a fetal biophysical profile," *Am. J. Obstet. Gynecol.*, vol. 136, pp. 787-795, 1980.

36. J.P. Phelan, "Tests of fetal well-being using the fetal heart rate," in *Fetal Monitoring; Physiology and Techniques of Antenatal and Intrapartum Assessment* , ed. J.A.D. Spencer, pp. 60-63, Castle House Publications Ltd, Kent, 1989.

37. F.A. Manning , "The Biophysical Profile ," in *Fetal Monitoring; Physiology and Techniques of Antenatal and Intrapartum Assessment* , ed. J.A.D. Spencer, pp. 73-84 , Castle House Publications Ltd, Kent, 1989.

38. J.B. Trimbos and M.J.N.C. Keirse, "Observer variability in assessment of antepartum cardiotocograms," *Br. J. Obstet. Gynaecol.*, vol. 85, pp. 900-906, 1978.

39. F.K. Lotgering, H.C.S. Wallenburg, and H.J.A. Schouten, "Interobserver and intraobserver variation in the assessment of antepartum cardiotocograms," *Am. J. Obstet. Gynecol.*, vol. 144, pp. 701-705, 1982.

40. H.E. Bassil and J.H. Dripps, "Real time processing and analysis of fetal phonocardiographic signals," *Clin. Phys. Physiol. Meas.*, vol. 10, Suppl. B, Institute of Physical Sciences in Medicine, UK, 1989.

41. J.H. Dripps, G.K. Manning, and X. Zhu, "A signal processing research facility and its application to the processing of fetal phonocardiographic signals for heart rate estimation," *Eur. J. Obstet. Gynecol. Reprod. Biol.*, vol. 23, pp. 281-8, 1986.

42. M.B. Priestly, *Spectral analysis and time series(Volume 1: Univariate series),* Academic Press , London , 1981.

43. W.C. Orr and H.J. Hoffman, "A 90-Min Cardiac Biorhythm: Methodology and Data Analysis Using Modified Periodograms and Complex Demodulation," *IEEE Trans. Biomed. Eng.*, vol. BME-21, no.2, pp. 130-143, 1974.

44. K.J. Dalton and A.J. Dawson, "Baseline: A Computer Method Of Calculating Baseline in Fetal Heart Rate Recordings," *Int. J. Bio-Medical Computing*, vol. 15, pp. 311-317, Elsevier Scientific Publishers Ireland Ltd, 1984.

45. J.W. Tukey, *An Introduction to the Frequency Analysis of time series,* Princeton Univ., Mathematics 596, Princeton, N.J., 1963. J.R. thompson and D.R. Brillinger, Ed.

46. A.V. Oppenheim, A.S. Willsky, and I.T. Young, *Signals and Systems,* Prentice-hall International, Inc., 1983.

47. S.M. Bozic, *Digital and Kalman Filtering,* Edward Arnold, 1979.

48. S. Kay and S. Marple, "Spectrum Analysis-A Modern Perspective," *Proc. IEEE.*, vol. 69, pp. 1380-1419, Nov. 1981.

49. M.R. Speigel, *Schaum's outline of theory and problems of Statistics* , McGraw-Hill Book Company, New York, 1972.

50. S-Y Yeh, A. Forsythe, and E.H. Hon, "Quantification of fetal heart beat-to-beat interval differences," *Obstet. & Gynaec.*, vol. 41, pp. 355-363, 1973.

51. B.K. Kirkwood , *Essentials of Medical Statistics,* Blackwell Scientific Publications, Oxford, 1988.

52. E.H. Hon, *An Introduction to Fetal Heart Rate Monitoring,* Harty Press, New Haven, Conneticut, 1969.

53. P. Rolfe , *Fetal and Neonatal Physiological Measurements* , Pitman Medical Press , Bath, UK , 1980.

54. M. James, *Classification Algorithms,* Collins, London, 1985.

55. W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, *Numerical Recipes in C, The Art of Scientific Computing,* Cambridge University Press, 1988.

56. D.G. Altman, *Practical Statistics for Medical Research,* Chapman and Hall, London, 1991.

57. G.B. Wetherill, *Intermediate Statistical Methods,* Chapman and Hall, New York, 1981.

58. A.C. Harvey, *Forecasting, structural Time Series Models and the Kalman filter,* Cambridge University Press, Cambridge, 1989.

59. D.C. Montgomery and L.A. Johnson, *Forecasting and Time Series Analysis,* McGraw-Hill, New York, 1976.

60. G.E.P. Box and M.G. Jenkins, *Time-Series Analysis: Forecasting and Control,* Holden-Day Series, 1976.

61. S. Taylor, *Modelling Financial Time Series* , John Wiley and Sons , 1986.

62. J. D. Cryer, *Time Series Analysis,* Duxbury Press, Boston, 1986.

63. B. L. Bowerman and R.T. O'Connell, *Time series Forecasting,* Duxbury Press, Boston, 1987.

64. A. Harvey, *The Econometric Analysis of Time Series,* Philip Allan Publishers, Exeter, 1988.

65. P.Z. Peebles, *Probability, Random Signals, and Random Signal Principles,* McGraw-Hill International Book Company, Tokyo, Japan, 1980.

66. R.E. Mortensen, *Random Signals and Systems,* Wiley Publication, New York, 1987.

67. E. Parzen, *Stochastic Processes,* Holden-Day,Inc., San Francisco, 1965.

68. J. Makhoul, "Linear Prediction: A Tutorial Review," *Proc. IEEE.*, vol. 63, pp. 561-580, 1975.

69. D. Graupe, J. Salahi, and D. Zhang, "Stochastic Analysis of Myoelectric Temporal Signatures for Multifunctional Single-Site Activation of Prostheses and Orthoses," *J. Biomed Eng.*, vol. 7, pp. 18-29, Butterworth & Co(Publishers) Ltd, Jan. 1985.

70. G. Hefftner, W. Zucchini, and G.G. Jaros, "The Electromygram(EMG) as a Control Signal for Functional Neuromuscular Stimulation-Part I: AR-Modelling as a Means of EMG Signature Descrimination," *IEEE Trans. Biom. Eng.*, vol. 35, pp. 230-237, Apr. 1988.

71. R.J. Triolo, D.H. Nash, and G.D. Moskowitz, "The Identification of Time series Models of Lower Extremity EMG for the Control of Prostheses Using Box-Jenkins Criteria," *IEEE Trans. Biom. Eng.*, vol. 35, No. 8, pp. 584-594, Aug. 1988.

72. M.S. Bartlett, "On the Theoretical Specification of Sampling Properties of Autocorrelated Time series," *J. Royal Stat. Soc.*, vol. B8, p. 27, 1946 .

73. R.B.Shumway , *Applied Statistical Time Series Analysis,* Prentice-Hall International, 1988.

74. W.W.S. Wei, *Time series Analysis Univariate and Multivariate Methods,* Addison-Wesley Publishing Company, Inc., Redwood City California , 1990.

75. K.S. Shanmugan and A.M. Breipohl, *Random Signals, Detection, Estimation and Data Analysis,* Wiley Publications, 1988.

76. T.W. Anderson, *The Statistical Analysis of Time Series,* Wiley Publications, New York, 1971.

77. E. Kreyszig, *Advanced Engineering Mathematics,* Wiley Publications, New York, 1979.

78. N. Levinson, "The Wiener RMS Error Crieterion in Filter Design and Prediction," *J. Math. Physics*, vol. 25, pp. 261-278, 1947.

79. J. Durbin, "The Fitting of Time Series Models," *Rev. Int. Inst. Statist.*, vol. 28, pp. 233-244, 1960.

80. M.H. Quenouille, "Approximate Tests of Correlation in Time-series," *J. Roy. Stat. Soc.*, vol. 11, pp. 68-84, 1949.

81. C. Chatfield , *The Analysis of Time Series: An Introduction,* Chapman and Hall , London , 1980.

82. P.J. Diggle, *Time Series: A Biostatistical Introduction,* Oxford Science Publications , New York, 1990.

83. A. Pankratz, *Forecasting with Univariate Box-Jenkins Models: Concepts and Cases,* Wiley Publications, New York, 1983.

84. H. Akaike, "Maximum Likelihood Identification of Gaussian auto-regressive moving average models," *Biometrika*, vol. 60, pp. 255-266, 1973.

85. R. Shibata, "Selection of the Order of an Autoregressive Model by Akaike's Information Criterion," *Biometrika*, vol. 63, pp. 117-126, 1976 .

86. J.M. Sneek, *Modelling Procedures for Economic Time Series,* Free University Press, Amsterdam, 1984.

87. E. Parzen, "Recent advances in time series modelling," *IEEE Trans. Auto. Contr.*, vol. 19, pp. 723-730, 1974.

88. W.S. Cleveland , "The inverse autocorrelations of a time series and their applications," *Technometrics*, vol. 14, pp. 277-298, 1972.

89. A.C. Harvey and G.D.A. Phillips, "Maximum Likelihood Estimation of Regression Models with Autoregressive Moving Average Disturbances," *Biometrika*, vol. 66, pp. 49-58, 1979 .

90. J.M. Mendel, *Optimal Seismic Deconvolution an estimation-based Approach,* Academic Press, 1983.

91. J.J. Shynk, "Adaptive IIR Filtering," *IEEE ASSP Magazine*, vol. 6, pp. 4-21, April 1989.

92. J.V. Candy, *Signal Processing The Model-based approach,* McGraw-Hill, 1986.

93. M.D. Srinath and P.K. Rajasekaran, *An Introduction To Statistical Signal Processing With Applications,* John Wiley&Sons, 1979.

94. R.K. Mehra, "Approaches to Adaptive Filtering," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 693-698, Oct. 1972.

95. B.D.O. Anderson and J.B. Moore, *Optimal Filtering,* prentice Hall, 1979.

96. R.H. Jones, "Maximum Likelihood Fitting of ARMA Models to Time Series With Missing Observations," *Technometrics*, vol. 22, pp. 389-395, 1980.

97. R.A. Thisted, *Elements of Statistical Computing Numerical Computation,* Chapman and Hall, 1988.

98. S.S. Rao, *Optimization theory and applications,* Wiley Eastern Ltd, New Delhi, 1978.

99. M.A. Wolfe, *Numerical Methods for unconstrained optimization(an introduction),* Van Nostrand Reinhold, 1978.

100. W.T. Vetterling, S.A. Teukolsky, W.H. Press, and B.P. Flannery, *Numerical Recipes, Example Book(C),* Cambridge University Press, 1989.

101. J.L. Kuester and J.H. Mize, *Optimization Techniques With Fortran,* McGraw-Hill Book Company, 1973.

102. P.E. Gill, W. Murray, and M.H. Wright, *Practical Optimization,* Academic Press, 1981 .

103. G.M. Ljung and G.E.P. Box, "On a Measure of lack of Fit in Time Series Models," *Biometrika*, vol. 65, pp. 67-72, 1978.

104. L.G. Godfrey, "Testing the Adequacy of a Time Series Model," *Biometrika*, vol. 66, pp. 67-72, 1979.

105. G.M. Jenkins and D.G. Watts, *Spectral Analysis and its applications,* Holden-Day , San Francisco, 1968.

PUBLICATIONS

[1]: M.A. Shariati, J.H., Dripps, "Stochastic time series analysis of fetal heart rate variability", IEEE Western Canada Conference on Telecommunication for Health Care, July 1990, Calgary, Published in the Proceedings of SPIE, Volume 1355, pp. 144-149, 1990.

[2]: M.A. Shariati, J.H. Dripps, K. Boddy, "Baseline Estimation Statistical and Numerical Analysis of Fetal Heart Rate", Proceedings of IV International Symposium On Biomedical Engineering, Peniscola - Spain, September, 1991.

Stochastic time series analysis of fetal heart rate variability

M.A. Shariati and J.H. Dripps

Department of Electrical Engineering,
King's Buildings, University of Edinburgh,
Edinburgh, EH9 3JL,UK.

ABSTRACT

Fetal Heart Rate(FHR) is one of the important features of fetal biophysical activity and its long term monitoring is used for the antepartum(period of pregnancy before labour) assessment of fetal well being. But as yet no successful method has been proposed to quantitatively represent variety of random, non-white patterns seen in FHR. Objective of this paper is to address this issue. In this study, the Box-Jenkins method of model identification and diagnostic checking was used on phonocardiographic derived FHR(averaged) time series. Models remained exclusively autoregressive(AR). Kalman filtering in conjunction with maximum likelihood estimation technique forms the parametric estimator. Diagnostics performed on the residuals indicated that a second order model may be adequate in capturing type of variability observed in 1 up to 2 min data windows of FHR. The scheme may be viewed as a means of data reduction of a highly redundant information source. This allows, a much more efficient transmission of FHR information from remote locations to places with facilities and expertise for closer analysis. The extracted parameters are aimed to reflect numerically the important FHR features. These are normally picked up visually by experts for their assessments. As a result long term FHR recorded during antepartum period, could then be screened quantitatively for detection of patterns considered normal or abnormal.

1. INTRODUCTION

Changes in FHR variability during antepartum period, can be used as an indicator of the well-being of the fetus. Normally FHR varies continually and the magnitude and frequency of these changes is referred to as the variability. The state of variability, baseline level, and presence of accelerations(detection of 5 per hour is considered a normal sign[1]), are the general type of information that are usually used as part of antepartum assessment[2] of fetal condition.

Normal tracings exhibit a baseline level [3] in the range 120 to 160 beats per minute(bpm) and an overall variability of 6-25 bpm. Variability less than 5 bpm is regarded as minimal or larger than 25 bpm as increased variability[4] . Detection of low variability(flattened pattern)[2] can be warning of fetal demise, although it has other associated meaning such as fetal prematurity, fetal sleep, or it can arise from the effect of drugs. Increased variability occasionally may indicate cardiac compensation following a reduction in blood flow.

In general there are two forms of variation: beat-to-beat changes(referred to as short term variability), and cyclic changes, with approximate periodicities of 10-20 seconds(referred to as long term variability). Short term variability can only be detected accurately from fetal electrocardiogram(FECG) obtained from a scalp electrode(direct monitoring), which is restricted to the phase of active labour and delivery[5] .

During antepartum period of monitoring only external modes of measurements of fetal heart signals are possible. In all of these modes, signals picked up to compute FHR are usually mixed up with noise, so to make the trace show a reasonable visual pattern, the FHR calculated, is of averaged form[5] . Thus the term averaged FHR is more of an appropriate name to describe FHR calculated from external or indirect methods.

FHR data used in our analysis was obtained from the external phonocardiographic method. In this approach fetal heart sounds is picked up from a sensitive microphone[6] placed on the maternal abdomen. The recorded heart sounds are then further processed to extract FHR [7] . Main advantage of this method is firstly its ease of application, and secondly it is non-invasive to both the mother and fetus, allowing long term recording of fetal heart data possible. The approach[7] produces an averaged FHR time-series, at a sampling rate of 0.5 Hz(averaging window of 2 sec duration). Due to relative short size of averaging window, long term variability is fully captured. Furthermore it is necessary to mention that obstetricians and midwifes when visually analyse FHR variability, base their judgement mainly on long term variations[5] .

Because of its relevance as an important screening tool[6] , many attempts has been made to quantify FHR variability, and almost all of them reduce to some form of estimation of standard deviation(or variance)[1,8] . This simple one-dimensional type of representation is inadequate, and a more complex approach is required for representing variety of random non-white patterns observed in averaged FHR time series.

We have focused our analysis to random long term variability, as shown in Fig. 1a and 1b. Objective is to quantify different patterns that variability exhibits, along with its respective baseline value. Together they offer valuable medical informations suitable for long term nonstress(NST)[5] monitoring of a fetus during the antepartum period. In our approach we are attempting to represent averaged FHR segments(1 min up to 2 min) in terms of five parameters. The first three are parameters of a stochastic process, a second order autoregressive(AR(2))[9] structure. These parameters are used to represent FHR variability. The other two are basic statistical quantities, mean(μ_p) and standard deviation(σ_p) of the segment under observations, used to represent the FHR baseline level and its relative statistical spread. The AR(2) model is defined by the equation

$$z(k) = \alpha_1 z(k-1) + \alpha_2 z(k-2) + a(k) \tag{1}$$

The parameters that we are interested in are the weights α_1 and α_2 and standard deviation σ_a of the white noise input $a(k)$ which drives the model. In the next section time-series analysis(time domain) will be discussed in detail.

2. TIME SERIES ANALYSIS METHODOLOGY

Data that are obtained from observations of a phenomenon over time are extremely common. The list of areas in which time series are observed and analyzed is endless. The purposes of time series analysis are generally two-fold: to understand or model the stochastic mechanism that gives rise to an observed series and to predict or forecast future values of a series based on the history of that series. Finding appropriate models for time series is not an easy task. There are three main steps in the model-building procedure, each of which may be used several times. These steps will be discussed next.

2.1 Model identification

To identify a stochastic model based on Box-Jenkins methodology[9] , the statistics of the time series itself must be examined, in effect allowing the data decide the structure of the time series representation. The models that arise could be pure autoregressive processes of order p(AR(p))[9] , or pure moving average processes of order q(MA(q)), or mixed autoregressive moving average(ARMA(p,q)). In general an AR model describes an all-pole filter driven by a white noise input(see eq. 1 for AR(2)), and a MA model describes an all-zero filter. Mixed ARMA models contain both AR and MA parts.

It can be mathematically proven that each of these has its own unique theoretical behaviour for autocorrelation(ACF) and partial autocorrelation(PACF) functions[9,10] , which results in several guidelines for model identification. Briefly the main ones are as follows:

1. For a MA(q) model, ACF has spikes at lags 1,2,..,q, and cuts off after lag q, and PACF dies down(damped exponential fashion with or without oscillation, or damped sine-wave fashion).

2. For an AR(p) model, ACF dies down(as PACF for MA(q)) and PACF has spikes at lags 1,2,..,p, and cuts off after lag p.

3. For a mixed model, both ACF and PACF die down.

4. ACF's which remain high for large lags indicate nonstationarity in the time series, which is usually removed by differencing the series.

The ACF and PACF can only be estimated from real data, so to be exact they should be addressed as sample partial autocorrelation function(SPAC) and sample autocorrelation function(SAC). In order to test the significance of the estimated correlation spikes, statistical tests of significance are necessary[10,11] . Fig. 3 and 4 show SAC and SPAC of two different boxed windows of averaged FHR shown in Fig. 1a and 1b respectively. Following the above criteria, an AR(2) model is appropriate for the given observed series. In choosing a model, one must attempt to adhere to the principle of *parsimony*; that is, the model used should require the smallest possible number of parameters that will adequately represent the data. The objective of this form of time series analysis is not to exactly reconstruct the time series itself as in analytical methods(high resolution spectral analysis)[12] , but to extract just the general features, which will help us to discriminate quantitatively between different patterns observed in the series.

2.2 Model fitting

Once a model has been(tentatively) specified, the parameters of the structure must be estimated. There are many ways in estimating the model parameters. But to come up with the right choice, the nature of the data must be considered. In case of averaged FHR, after the removal of its low order trends, it is still a time varying process, exhibiting nonstationarity(second moment). Maximum likelihood estimation technique in conjunction with Kalman filtering[11,13] has the capacity to deal(to a certain degree)

with the nonstationary time series. Although it is a relatively difficult approach to the problem; it offers consistency, stability and unbiased estimation[13,14] , especially for pure AR models[15] .

Furthermore the advantage of the method of maximum likelihood is that all of the information available in the data is used(rather than just the first and second moments). As a result it is more efficient and less dependable on large samples. In addition the use of Kalman filtering allows the estimator to be readily adapted to series observed with noise, or missing data[11,16] .

Briefly, the maximum likelihood estimate(MLE) $\hat{\theta}$ is the value of θ that maximises log-likelihood L for a particular set of observations z. L is defined as,

$$L = L\{\theta|z\} = \ln l(\theta|z) \propto \ln p(z|\theta), \tag{2}$$

where l is likelihood of θ, given the observation z, and p is the conditional joint probability density function. The likelihood is defined to be proportional to the value of probability density function of the observations given the parameters. The logarithm of l is a monotonic transformation of l. Now obtaining an expression for $L(\theta|z)$ is not a trivial matter, since the observations are not statistically independent. The residuals(innovations) generated from a Kalman filter(based on the identified model)[13] can be used to solve the problem of statistical dependency in the time series. Basically they are used to provide the information needed, to define the above log-likelihood function. The expression derived is then maximised with respect to the model parameters. When the optimisation routine is over, the residuals should resemble a white sequence(that is if the model is adequate). For further detail please refer to references [13] , and[11].

2.3 Model diagnostics

Once a model has been identified and the parameters estimated, diagnostic checks must be applied to the model to determine its adequacy. These checks are normally applied to the residuals(prediction errors) of an estimation process. Residuals are basically error terms representing the difference between raw data($z(k)$;k is discrete time index) and their corresponding model estimates($\hat{z}(k)$). The model will only be adequate, when its corresponding residuals are white. A first step in this checking process is a visual inspection of a plot of the residuals. For adequacy they should exhibit random white characteristics. Checks are then applied to the SAC and SPAC of the residuals time series. If the model is correct, then their should not be any significant correlation spikes from the two functions. The spikes should be distributed, approximately normally about zero. They should lie within plus or minus two standard deviation of $2/\sqrt{N}$, N being the number of residuals[17] .

A more formal test, known as the *portmanteau lack of fit test*, considers the sample autocorrelations of the residuals as a whole, rather than individually, when used to indicate the adequacy of a model. The test calculates a value Q according to the formula[18]

$$Q = N(N+2)\sum_{k=1}^{K}\frac{1}{N-k}\hat{r}_k^2 \tag{3}$$

where N is the number of residuals and K is the number of lags and $\hat{r}_k$ is sample autocorrelation estimate at lag k(common values for K are 9 or 17). If the model is correct, Q has a chi-square distribution with (K-p-q) degree of freedom. Fitting an erroneous model causes the residuals to be more related, thus Q inflates. In general a model can be rejected if its Q statistics is greater than the *chi −square critical value* $\chi^2_{0.05}[K-p-q]$.

3. RESULTS

Before the variability could be modeled, mean and low order trends must first be removed. The low order trends can be easily removed by a quadratic regression fit[11] . The deviations from the regression line can then be used as input for the variability analysis. A model order must be chosen which will generally produce a reasonably well-fitted model. Through applications of all the steps described previously, the adequacy of an AR(2) model for modeling long term variability couldn't be rejected. Fig. 2a and 2b show synthetic data generated from an AR(2) model. These are produced from the estimated parameters extracted from the boxed segments of FHR shown in Fig. 1a and 1b. The visual resemblance between the variability in the real data and its corresponding synthetic version is apparent. The SAC and SPAC of these boxed segments are shown in Fig 3 and 4, clearly identifying an AR(2) process. Furthermore Fig. 6 and 7 show the SAC and SPAC of the model residuals(Fig. 5), both showing adequate fit of the model. The Q statistics were also low in value, well below the chi-square critical value.

In these two example a data window of 2 min(60 samples) is represented in terms of five parameters. The first three α_1, α_2 and σ_a of the AR(2) model reflects the dynamics of long term heart rate variability. The other two, are scalar quantities, μ_p and σ_p representing the baseline level and the statistical spread(over the baseline) in the segment under observation.

4. CONCLUSION AND DISCUSSION

The variability(long term) observed in externally monitored FHR(averaged) is stochastically well-fit to a univariate stochastic process(AR(2)). The parameters estimated represents fairly the dynamics of FHR variability(long term). The synthetic data generated from the parameter estimates, bear a close visual resemblance to the original data, providing additional confirmation that the data are adequately represented. By contrast, other scalar measures of long term variability cannot distinguish among the variety of patterns which are seen clinically. Thus the AR-parameters could be used to discriminate between variability patterns. The technique can be looked upon as means of data reduction. This is of value, when we are dealing with a highly redundant information source, such as long term recordings of FHR.

The future objective is to use this type of time series analysis approach as part of a complete user friendly FHR screening system. Such a system could be installed in any remote location. The features extracted could then be send over a telephone link(or any other form of transmission) to places with facilities and expertise, for careful analysis and screening.

Finally this technique can be easily adapted to other important biomedical time series, such as adult heart rate dynamics, breathing rhythm, gait analysis and electroencephalography.

5. ACKNOWLEDGEMENTS

Authors would like to thank Mr H.E. Bassil for the provision of externally recorded FHR information, used in their time series analysis. Also would like to add that, this work was partially sponsored by the European Advanced Informatics in Medicine(AIM) Initiative Through "Telemedicine" project(A-1032).

6. REFERENCES

References

1. Devoe, L.D., "The Fetal Biophysical Profile: Antepartum Assessment Using a Programmed Microcomputer," *J. Clinicl. Eng.*, vol. 11, pp. 285-289, Quest & Co, Jul-Aug, 1986.
2. vanGeijn, H.P. , "Fetal monitoring-present and future: The evaluation of fetal heart rate patterns," *Europ. Jnl. of Obst & Gyn.*, vol. 24, pp. 117-119, 1987.
3. Trierweiler, M.W., Freeman, R.K., and James, J., and Hon, E.H., "An Introduction to Fetal Heart Rate Monitoring," *Amer. J. Obstet. Gynecol.*, vol. 125, pp. 618-623, Harty Press, New Haven, Conneticut, 1969.
4. Fleming, J. E. E., *Fetal Monitoring*, pp. 332-339, Pergamon Press, Glasgow, UK, 1984. From 'Physics in Medicine & biology Encyclopedia;T.F. McAinsh"
5. Kunzel, W., *Fetal Heart Rate Monitoring*, Springer-Verlag, Heidelberg, 1985.
6. Carter, M.C., "Fetal Monitoring," *J. Biomed. Eng.*, vol. Vol. 10, pp. 527-532, Butterworth & Co, 1988.
7. Bassil, H.E. and Dripps, J.H., "Real time processing and analysis of fetal phonocardiographic signals," *Clin. Phys. Physiol. Meas.*, vol. 10, Suppl. B, Institute of Physical Sciences in Medicine, UK, 1989.
8. Angel, E.S., Fox, H.E., and Titlebaum, E.L., "Digital Filtering and Fetal Heart Rate Variability," *Comp. & Biom. Research*, vol. 12, pp. 167-180, 1979.
9. Box, G.E.P. and Jenkins, M.G., *Time-Series Analysis: Forecasting and Control*, Holden-Day Series, 1976.
10. Bowerman, B. L. and O'Connell, R.T., *Time series Forecasting*, Duxbury Press, Boston, 1987.
11. Cryer, J. D., *Time Series Analysis*, Duxbury Press, Boston, 1986.
12. Kay, S. and Marple, S., "Spectrum Analysis-A Modern Perspective," *Proc. IEEE.*, vol. 69, pp. 1380-1419, Nov. 1981.
13. Mendel, J.M., *Optimal Seismic Deconvolution an estimation-based Approach*, Academic Press, 1983.
14. Shynk, J.J., "Adaptive IIR Filtering," *IEEE ASSP Magazine*, vol. 6, pp. 4-21, April 1989.
15. Bennett, R.J., *Spatial Time Series: Analysis-Forecasting-Control*, Page Bros (Norwich) Ltd, London, 1979.
16. Thisted, R.A., *Elements of Statistical Computing Numerical Computation*, Chapman and Hall, 1988.
17. Quenouille, M.H., "Approximate Tests of Correlation in Time-series," *J. Roy. Stat. Soc.*, vol. 11, pp. 68-84, 1949.
18. Ljung, G.M. and Box, G.E.P., "On a Measure of lack of Fit in Time Series Models," *Biometrika*, vol. 65, pp. 67-72, 1978.

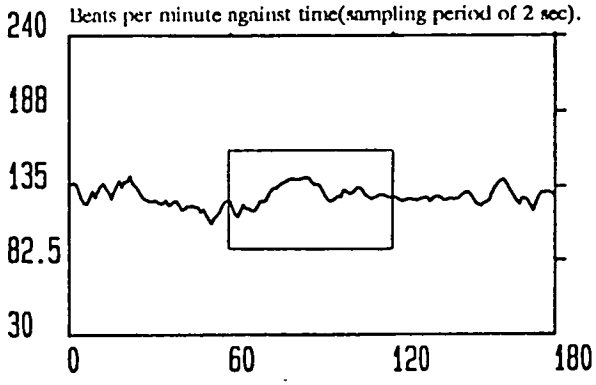


Fig. 1a. Averaged FHR(indirect phonocardiography).

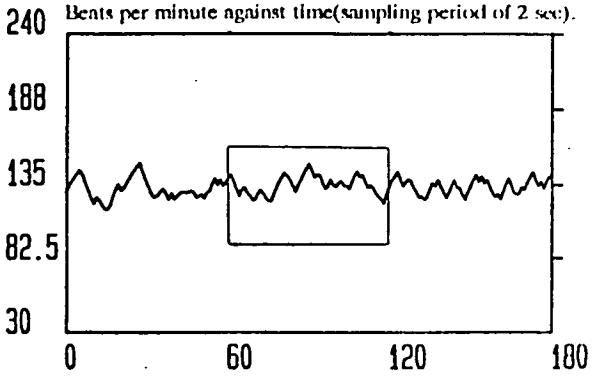


Fig. 1b. Averaged FHR(indirect phonocardiography).

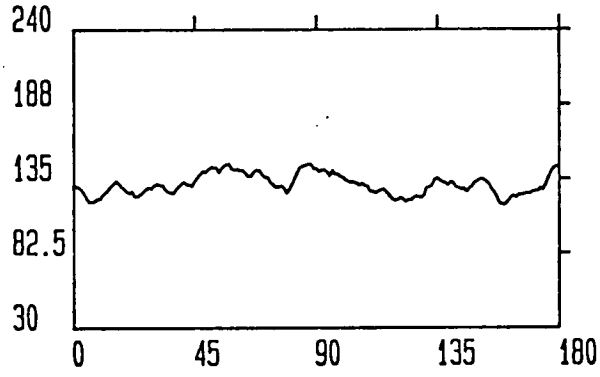


Fig. 2a. Synthetic FHR data, generated from the parameters extracted from the boxed segment shown in fig. 1a.

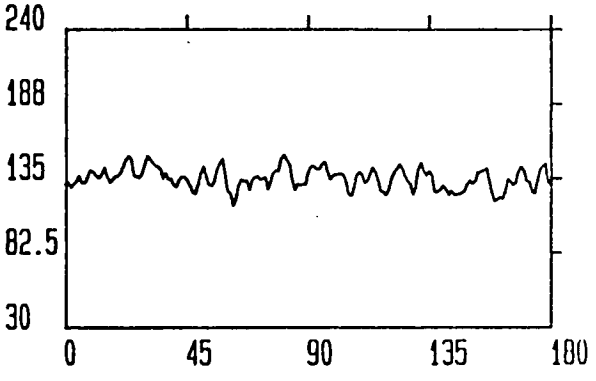


Fig. 2b. Synthetic FHR data, generated from the parameters extracted from the boxed segment shown in fig. 1b.

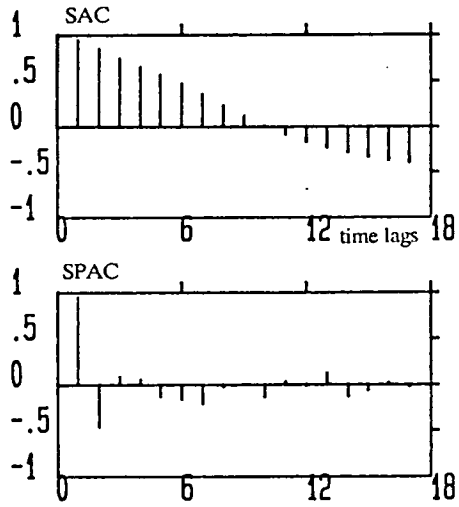


Fig. 3. SAC and SPAC of boxed segment in fig. 1a; indicating an AR(2) process.

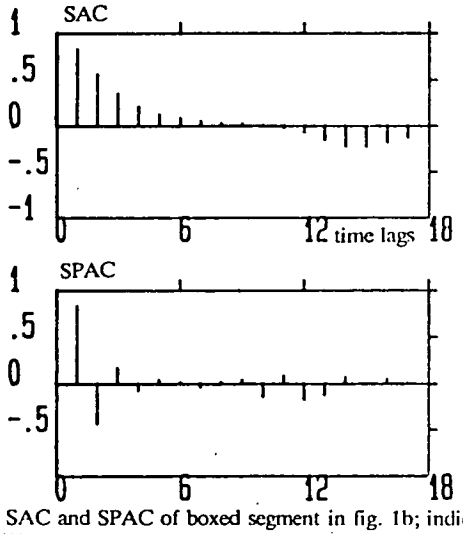


Fig. 4. SAC and SPAC of boxed segment in fig. 1b; indicating an AR(2) process.

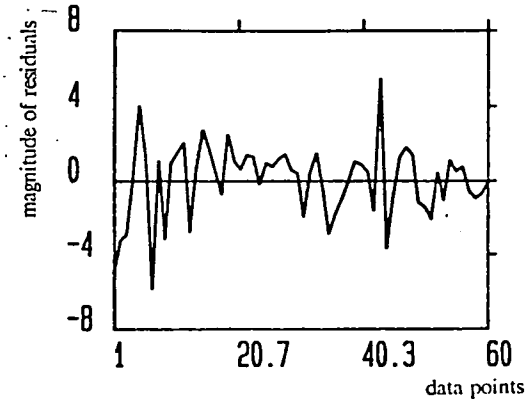


Fig. 5a. Residuals left from fitting an AR(2) model to the boxed segment in fig. 1a.

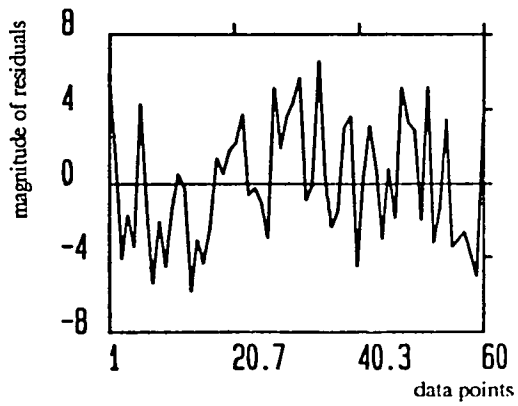


Fig. 5b. Residuals left from fitting an AR(2) model to the boxed segment in fig. 1b.

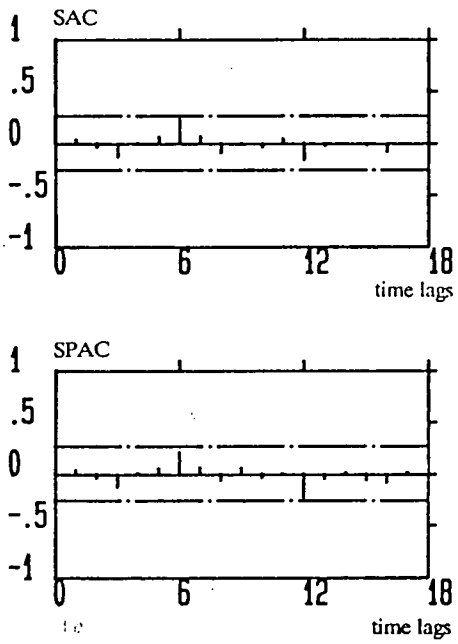


Fig. 6. SAC and SPAC of residuals shown in fig. 5a, indicating a white sequence. The dot-dashed lines are confidence limits($\pm 2\sigma$) for the insignificant spikes.

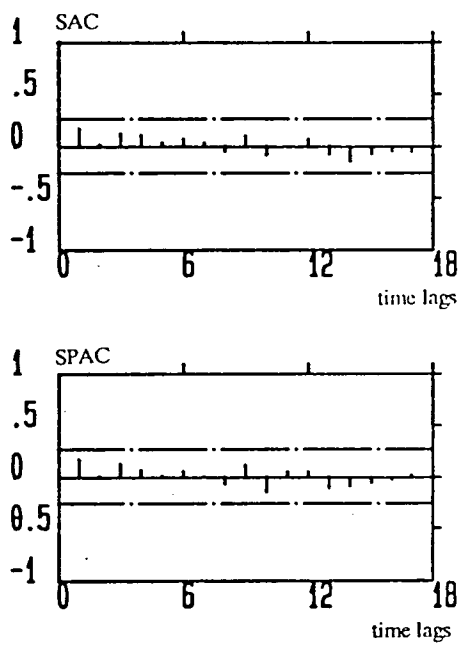


Fig. 7. SAC and SPAC of residuals shown in fig. 5b, indicating a white sequence.

BASELINE ESTIMATION STATISTICAL AND NUMERICAL ANALYSIS OF FETAL HEART RATE

M.A. Shariati †,J.H. Dripps †,K. Boddy ‡

†Department of Electrical Engineering, University of Edinburgh.

‡Department of Obstetric and Gynaecology, University of Edinburgh.

Abstract-Biomedical time series such as fetal heart rate(FHR), exhibit varying shapes and patterns. Almost all will in some way or another exhibit non-stationarity in their first(mean) and second moments(variance). In this paper a first order bi-directional digital autoregressive filtering technique will be introduced for unbiased estimation of averaged FHR slow underlying rhythm(baseline or trend), a technique which also would be most useful and applicable to other biomedical data. Successful derivation of baseline information, paved the way towards a correct statistical analysis approach of long term FHR variability, and derivation of a successful rule based numerical detection strategy for FHR accelerations which is an important medical screening component for assessment of fetal condition.

1. INTRODUCTION

Biomedical data often are non-stationary. They usually have slow underlying rhythms(e.g. adult or fetal heart rate data) also known as baseline or trend(time varying first moment), and continuously changing level of fluctuations over a changing trend(time varying second moment).

Proper unbiased estimation of baseline is of paramount importance and value. In the case of FHR, the estimation of its baseline, is an essential prerequisite for proper quantitative analysis of its characteristics and patterns.

In general there are two methods of estimating and removing the baseline. One way is to fit to the data, on a contiguous block basis, different deterministic models, such as the simplest linear trend model, or nonlinear polynomial models. However when dealing with long hours of measurements, this approach for over-all estimation of and removal of the baseline will result in unquantifiable statistical bias in the estimated trend and in the detrended data.

This problem can be remedied, if some form of linear digital filtering is used to estimate and remove the baseline. The approach reported here is based on first order linear bi-directional autoregressive filtering. The theory behind the approach was introduced by J.W. Tukey in 1963[1]. It was successfully applied to estimate adult human heart rate by Orr [2], and later on to estimate baseline of ultrasonically derived averaged FHR by Dalton[3]. Because of the bi-directional structure of this filtering technique, phase shift is completely eliminated. In the case of FHR data, the estimated baseline can then be subtracted from the original data to give the mean stationary(detrended FHR) data, which can then be used for quantitative statistical analysis of long term FHR variability.

In summary the bi-directional filtering technique will be introduced first. The detrimental effect of the baseline on the standard deviation statistic, which is usually used as a coefficient to measure the spread of irregularities of biomedical time series will be clearly shown. Finally the computed rate of change of the baseline made it possible to numerically detect accelerations. The theory behind this detection method will also be described in detail.

2. ESTIMATION OF BASELINE

Baseline or trend of a time series can be calculated by a first order purely recursive autoregressive filter applied to each time series twice: first in the forward direction and then in the reverse direction.

Equations (1) and (2) represent the two autoregressions, plus a final equation (3) that utilizes the output from the two stage autoregressive filter z(k), and the original time series x(k):

$$y(k) = \rho y(k-1) + (1-\rho)x(k) \quad (1)$$

$$z(k) = \rho z(k+1) + (1-\rho)y(k) \quad (2)$$

$$u(k) = z(k) - (1-b)x(k) \quad (3)$$

where

$$b = \frac{2\rho}{(1+\rho^2)}, \text{ and } 0.70 < \rho < 0.95$$

Equation (1) represents the time forward first order autoregression on the input FHR time series x(k), which yields the output series y(k). Equation (2) represents the second autoregression, which is applied to the output series y(k) from (1). The time index increments in the opposite direction in (2). The final output series from this autoregressive scheme is u(k).

The transfer function for this overall procedure for determining the trend is reduced to:

$$H_T(w) = (1-b)\frac{bcos(w)}{1-bcos(w)}, \; where \; (0 \leq w \leq \pi) \quad (4)$$

Since the transfer function is real valued, there has been no overall phase shift introduced by the bidirectional autoregressive filtering scheme. Hence, the phase shift introduced by (1) is exactly compensated by the equal but opposite-in-sign phase shift resulting from (2).

The value of ρ in equations (1) and (2) determines the 3 dB cut-off point of this bi-directional autoregressive filter. In general as ρ decreases, the bandwidth of the filter increases.

A value for ρ must be chosen such that the detrended FHR time series will not exhibit any non-stationarity in the form of a pronounced trend which will result in constant variation of the overall mean of the series. Through numerous trials on phonocardiographic derived averaged FHR[4] a value of 0.9 was found appropriate for ρ. Using this value resulted in a good estimate of the overall baseline of the FHR data which also includes the long-lasting accelerations and decelerations as part of the estimated trend. This property is exploited later on for the purpose of detecting accelerations.

The coloured residuals or the detrended FHR series is obtained by subtracting the smooth trend u(k) from the original time series x(k). Figures 1 to 3 show how three 24 minute blocks of averaged FHR(corresponding to 720 samples) have been detrended, acquiring stationarity in the first moment.

3. STATISTICAL ANALYSIS OF FHR VARIABILITY

Trend has an enormous influence on the value of the standard deviation statistic, effectively overshadowing the contribution that FHR variability makes, rendering it as a useless quantity for representing true variability over the baseline, when pronounced first moment non-stationarity is present. This is clearly demonstrated by the computed statistics(shown in table 1.) from the data segments shown in figures 1 to 3. The much higher standard deviations computed from the original data, is due to the presence of strong mean non-stationarity in those segments. Whereas the computed standard deviation from the detrended data, reflects the true level of FHR variability.

However the detrended FHR data still is nonstationarity in the second moment, i.e its variance continuously changes with time. Therefore in order to capture correctly the development with time of long term FHR variability, shortest possible observational epoch should be taken.

Statistically speaking for an unbiased second order moment estimation, the number of samples in an observation window must be at least 30 or higher if possible[5]. Any sampling block equal to 30 or above is considered statistically as being a *large sampling* block.

Therefore the shortest observational epoch which could be used for standard deviation estimation were limited to one minute only, exactly equivalent to a sampling block size of 30. This size of the window,also was short enough to have local statistical stationarity in the second moment.

Effectively one can look upon the one minute block by block standard deviation estimate of detrended data, as a one-dimensional non-parametric way of representing quantitatively evolutionary variance of long term FHR variability.

4. NUMERICAL DETECTION OF ACCELERATIONS

Detection of accelerations in the FHR-trace is seen as a positive sign in regard to fetal health. Obstetricians are satisfied as a sign of normality to detect four or more genuine accelerations per hour(associated to fetal movement) from FHR recordings made during the antepartum period.[6, 7].

In general accelerations are picked up visually. With the use of the unbiased estimated baseline discussed previously and a rule based routine that have been devised, they can now be detected by a computer numerically.

The routine mainly bases its detection of accelerations on the general rule that an acceleration is *genuine* if and only if there has been a rise of at least 15 beats per minute, lasting a minimum of 15 seconds from the time the FHR leaves the baseline until it returns to a lower steady state baseline during a 10 minute moving window [7, 8].

The rule-based routine devised makes use of the rate of change of the baseline and the original recorded averaged FHR time series to achieve its objective.

The rate of change of the baseline is approximately computed via the following equation:-

$$ROAD[k] = \frac{BAS[k+1]-BAS[k]}{T} \qquad (5)$$

where ROAD[k] represents the rate of accelerations and decelerations at k-th sample and BAS[k] the estimated baseline at k-th sample and T the sampling period.

Figure 4 is an example showing the original FHR and its estimated baseline and the rate of change of the baseline. It is clear for each acceleration, there is a distinct common rate of change of the baseline pattern.

First of all for an acceleration from the steady state baseline where there is no trend and the approximate rate of change is zero, there is always a deceleration where the rate falls back to a steady state baseline level, and the rate of change is approximately equal to zero again.

At the same time when the accelerations reach a steady state top average value, the rate of change of the baseline during this period will also be approximately zero. These rules along with the rule of what makes a *genuine* acceleration generally make up the structure of the devised rule based acceleration detector.

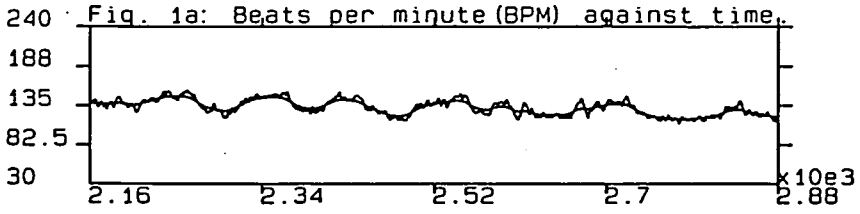
Referring back to figure 4, for each acceleration, its corresponding rate of change pattern, start from a value of zero, passes a heuristically found positive threshold level(+TL), peaks, returns to zero and goes negative. It then passes a heuristically found negative threshold level(−TL), peaks in the negative direction and returns back to zero again, to a point associated to completion of a possible acceleration, where the FHR baseline is in steady state form again.

Finally if it also satisfies the *genuine* acceleration condition mentioned earlier, then an acceleration is found, and its vital numerical representation(which includes starting, peak and ending point along with their FHR values) is recorded, as shown diagrammatically by the vertical lines drawn in figure 4.

In summary The Routine numerically scans forward two data sets(1. the original FHR data, 2. the rate of change of the baseline), looks for a possible acceleration, checks if it is a genuine acceleration, if it is, it proceeds in scanning back and forth the vicinity of the detected acceleration, to extract the relevant numerical information.

5. CONCLUSION

A successful trend estimation and removal technique has been introduced, which can also be applied to any other biomedical time series.



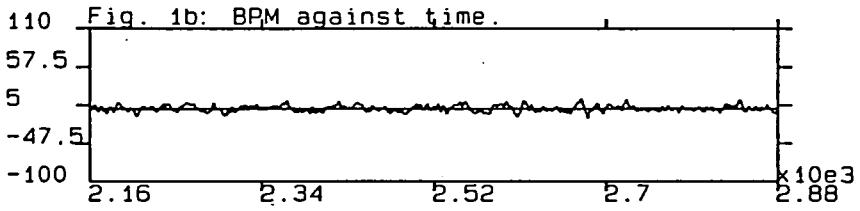
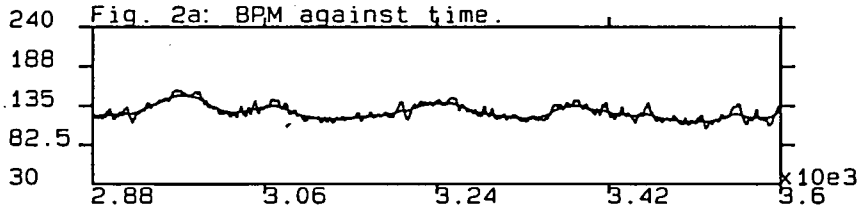


Fig. 1: Part (a) represents 24 minutes of averaged FHR data, along with its estimated baseline, and part (b) shows its detrended mean stationary version.



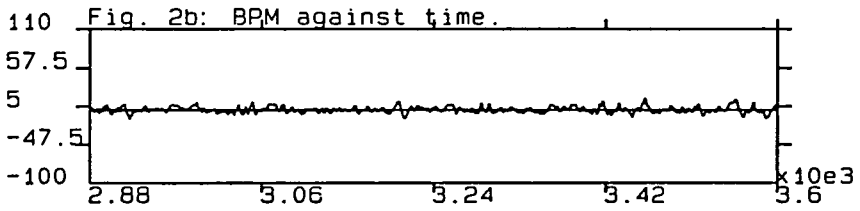
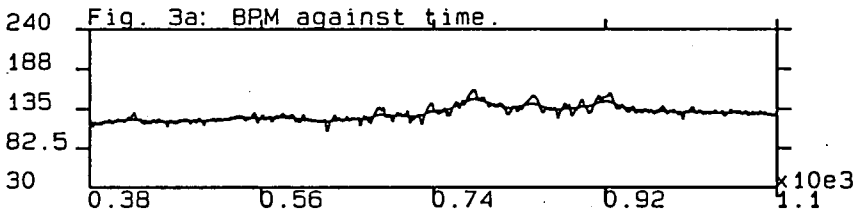


Fig. 2: Part (a) represents 24 minutes of averaged FHR data, along with its estimated baseline, and part (b) shows its detrended mean stationary version.



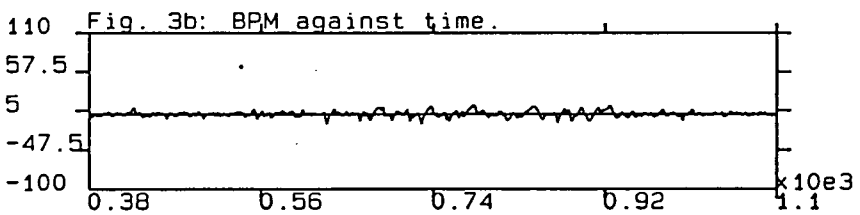


Fig. 3: Part (a) represents 24 minutes of averaged FHR data, along with its estimated baseline, and part (b) shows its detrended mean stationary version.

	Mean(BPM)		Variance (BPM^2)		Standard Deviation (BPM)	
Recording	Averaged FHR	Estimated Trend	Averaged FHR	Detrended FHR	Averaged FHR	Detrended FHR
Fig. 1.	132.483	132.173	113.944	17.507	10.674	4.184
Fig. 2.	126.442	126.228	107.633	17.181	10.375	4.145
Fig. 3.	128.838	128.680	91.366	12.614	9.558	3.552

Table 1. Showing Mean of averaged FHR(Indirect Phonocardiographic) and its corresponding Trend. Also showing the Variance and Standard deviation of FHR and its corresponing detrended version. The duration of the three recordings shown were all 24 minutes.

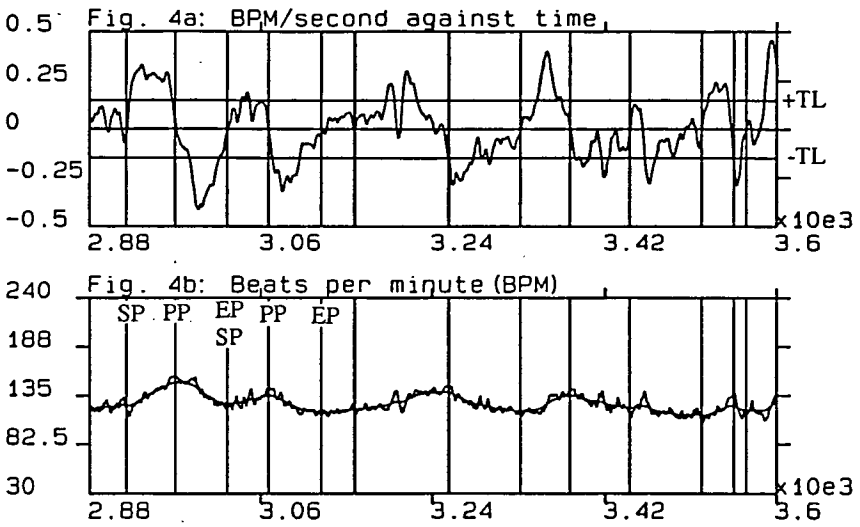


Fig. 4: Part (a) represents the rate of change of the baseline, shown in part (b) of the averaged FHR. The corresponding vertical lines drawn in (a) and (b) mark the points associated to starting(SP), peak(PP) and ending(EP) of accelerations detected, by the rule based routine devised.

The effect of mean nonstationarity or the FHR trend on the standard deviation statistic, which is the general scalar statistic used for the variability analysis, has been investigated. As a result it was concluded that in order to represent the genuine spread of the variability over the baseline, through the use of the scalar standard deviation statistic(or indeed any related statistic), it is vitally important to use detrended FHR data instead of the original.

Finally the other vital component of FHR, the accelerations, which are scanned for, visually by the medics, can now be accurately detected through the ruled based structure that have been devised. The routine is non-exhaustive, unlike the human expert and can accurately scan hours of antepartum recorded FHR data for detection of accelerations.

REFERENCES

1. Tukey, J.W., *An Introduction to the Frequency Analysis of time series,* Princeton Univ., Mathematics 596, Princeton, N.J., 1963. J.R. thompson and D.R. Brillinger, Ed.
2. Orr, W.C. and Hoffman, H.J., "A 90-Min Cardiac Biorhythm: Methodology and Data Analysis Using Modified Periodograms and Complex Demodulation," *IEEE Trans. Biomed. Eng.*, vol. BME-21, no.2, pp. pp. 130-143, 1974.
3. Dalton, K.J. and Dawson, A.J., "Baseline: A Computer Method Of Calculating Baseline in Fetal Heart Rate Recordings," *Int. J. Bio-Medical Computing*, vol. 15, pp. 311-317, Elsevier Scientific Publishers Ireland Ltd, 1984.
4. Bassil, H.E. and Dripps, J.H., "Real time processing and analysis of fetal phonocardiographic signals," *Clin. Phys. Physiol. Meas.*, vol. 10, Suppl. B, Institute of Physical Sciences in Medicine, UK, 1989.
5. Speigel, M.R., *Schaum's outline of theory and problems of Statistics* , McGraw-Hill Book Company, New York, 1972.
6. Devoe, L.D., "The Fetal Biophysical Profile: Antepartum Assessment Using a Programmed Microcomputer," *J. Clinicl. Eng.*, vol. 11, pp. 285-289, Quest & Co, Jul-Aug, 1986.
7. Phelan, J.P., "Tests of fetal well-being using the fetal heart rate," in *Fetal Monitoring; Physiology and Techniques of Antenatal and Intrapartum Assessment* , ed. J.A.D. Spencer, pp. 60-63, Castle House Publications Ltd, Kent, 1989.
8. Kunzel, W., *Fetal Heart Rate Monitoring,* Springer-Verlag, Heidelberg, 1985.

M.A. Shariati,
Department of Electrical Engineering,
University of Edinburgh, Kings Buildings,
Mayfield Road, Edinburgh,
EH9 3JL, Scotland, UK.
(Tel: 031-229 2517)